

FR2771103

Publication Title:

Isolated DNA encoding sodium channel of the nervous system

Abstract:

Isolated DNA sequences (I) of 5908 bp (1, from rat) and 856 bp (3, human). Independent claims are also included for the following: (a) expression vector containing (I); (b) host cell transformed with this vector; (c) recombinant DNA (Ia) including a sequence derived from (I); (d) sodium channel protein (II) encoded by (I) or its allelic variants; (e) use of (II) for identifying inhibitors of sodium channel proteins that are resistant to tetrodotoxin (TTX); (f) mono- or poly-clonal antibodies (Ab) directed against (II); and (g) diagnostic kits containing (Ia) able to hybridize specifically to (I).

Data supplied from the esp@cenet database - <http://ep.espacenet.com>

BEST AVAILABLE COPY

⑨ RÉPUBLIQUE FRANÇAISE
INSTITUT NATIONAL
DE LA PROPRIÉTÉ INDUSTRIELLE
PARIS

⑪ N° de publication :
(à n'utiliser que pour les
commandes de reproduction)

2 771 103

⑫ N° d'enregistrement national : **98 14551**

⑮ Int Cl⁶ : C 12 N 15/12, C 12 N 5/10, 15/85, C 07 K 14/47, 16/18, C 12 Q 1/68

⑫

DEMANDE DE BREVET D'INVENTION

A1

⑯ Date de dépôt : 19.11.98.

⑰ Priorité : 20.11.97 US 00066225.

⑴ Demandeur(s) : F. HOFFMANN-LA ROCHE AG
AKTIENGESELLSCHAFT — CH.

⑲ Date de mise à la disposition du public de la demande : 21.05.99 Bulletin 99/20.

⑵ Liste des documents cités dans le rapport de recherche préliminaire : *Ce dernier n'a pas été établi à la date de publication de la demande.*

⑵ Inventeur(s) : DIETRICH PAUL SHARTZER, FISH LINDA MARIE, KHARE REENA, RABERT DOUGLAS KENNETH et SANGAMESWARAN LAKSHMI.

⑶ Références à d'autres documents nationaux apparentés :

⑶ Titulaire(s) :

⑷ Mandataire(s) : REGIMBEAU.

⑸ SEQUENCE D'ADN CODANT POUR UNE PROTEINE DE CANAL SODIQUE, SA PRODUCTION ET SON UTILISATION.

⑹ L'invention décrit l'isolement d'une nouvelle séquence d'acide nucléique codant pour un canal sodique mammalien dépendant du potentiel, de préférence résistant à la tétrodoxine. Elle décrit également des produits polypeptidiques d'expression recombinante de ces séquences, des vecteurs d'expression comprenant la séquence d'ADN et des cellules hôtes transformées par ces vecteurs d'expression. D'autres aspects de la présente invention sont des peptides dont les séquences sont basées sur les séquences d'acides aminés déduites de ces séquences d'ADN, des anticorps spécifiques de ces protéines et de ces peptides, des procédés pour la détection et la détermination quantitative de telles protéines et d'acides nucléiques correspondants. Un autre aspect de l'invention est l'utilisation de ce canal sodique dépendant du potentiel, de préférence résistant à la tétrodoxine, comme cible thérapeutique pour des composés.

FR 2 771 103 - A1



Séquence d'ADN codant pour une protéine de canal sodique, sa production et son utilisation

La présente invention concerne d'une façon générale des protéines de canaux sodiques et plus particulièrement
5 une nouvelle séquence d'acide nucléique codant pour une sous-unité α mammalienne d'une protéine de canal sodique de tissu nerveux, dépendant du potentiel, de préférence résistante à la tétródotoxine. La présente invention concerne en outre sa production par la technique de recombinaison.
10

L'unité de base d'information transmise d'une partie du système nerveux à une autre est un seul potentiel d'action ou influx nerveux. La "ligne de transmission" pour ces influx est l'axone ou fibre nerveuse. L'excitabilité électrique de la membrane du nerf s'est avérée
15 dépendre du système de perméabilité ionique sensible à la tension de la membrane, qui lui permet d'utiliser l'énergie emmagasinée dans des gradients de concentration ionique. L'activité électrique du nerf est déclenchée par
20 une dépolarisation de la membrane, qui ouvre à travers la membrane des canaux qui sont hautement sélectifs pour les ions sodium, qui sont ensuite entraînés vers l'intérieur par le gradient électrochimique. Parmi les nombreux canaux ioniques, le canal sodique dépendant du potentiel ou sensible au potentiel est l'un des plus étudiés. Il s'agit
25

d'une protéine transmembranaire qui est essentielle pour la production de potentiels d'action dans des cellules excitables. Un excellent article de synthèse sur les canaux ioniques est donné par Catterall, *TINS* 16(12), 500-506 (1993).

Les ADNc pour plusieurs canaux de Na⁺ ont été clonés et séquencés. Numa et coll., *Annals of the New York Academy of Sciences* 479, 338-355 (1986), décrivent un ADNc provenant de l'organe électrique de l'anguille et deux ADNc différents provenant du cerveau de rat. Rogart, US-A-5 380 836, décrit un ADNc provenant de tissu cardiaque de rat; voir également Rogart et coll., *Proc. Natl. Acad. Sci.* 86, 8170-8174 (1989). Les séquences de PN1 et de ses orthologues chez l'homme (hNE) et le lapin (Na⁺) ont été publiées [voir par exemple Klugbauer et coll., *EMBO J.* 14, 1084-1090 (1995) et Belcher et coll., *Proc. Natl. Acad. Sci. USA* 923, 11034-11038 (1995)]. La séquence de PN1 de rat cloné à partir de ganglions de la racine dorsale (GRD) et son expression de fonction ont été décrites [voir par exemple Sangameswaran et coll., *J. Biol. Chem.* 272, 14805-14809 (1997)]. D'autres canaux sodiques clonés comprennent les types I et II [Noda et coll., *Nature* 320, 188-192 (1986)], IIa [Auld et coll., *Neuron* 1, 449-461 (1988)] et III [Kayano et coll., *FEBS Lett.* 228, 187-194 (1988)] de cerveau de rat, les canaux sodiques de muscle squelettique (SkM1) [Trimmer et coll., *Neuron* 3, 33-49 (1989)], NaCh6 de rat [Schaller et coll., *J. Neurosci.* 15, 3231-3242 (1995)], le canal sodique type III de nerf périphérique (rPN3) [Sangameswaran et coll., *J. Biol. Chem.* 271, 5953-5956 (1996), également dénommé SNS, Akopian et coll., *Nature* 379, 257-262 (1996)], le canal atypique de rat [Felipe et coll., *J. Biol. Chem.* 269, 30125-30131 (1994)] et le canal sodique glial de rat [Akopian et coll., *FEBS Lett.* 400, 183-187 (1997)].

Ces études ont montré que la séquence d'acides aminés du canal sodique avait été conservée pendant une longue période d'évolution. Ces études ont également révélé que le canal était un seul polypeptide contenant quatre
5 séquences répétées internes, ou domaines homologues (domaines I-IV) ayant des séquences d'acides aminés similaires. Chaque domaine se replie en six segments transmembranaires hélicoïdaux et prédits: cinq sont des segments hydrophobes et un est un segment hautement chargé,
10 comportant de nombreux résidus lysine et arginine chargés positivement. Ce segment hautement chargé est le quatrième segment transmembranaire dans chaque domaine (le segment S4) et est susceptible d'être impliqué dans l'activation par le potentiel. Les chaînes latérales chargées
15 positivement sur le segment S4 sont susceptibles de s'apparier avec les chaînes latérales chargées négativement sur les cinq autres segments, de sorte que la dépolarisation de la membrane peut déplacer la position d'une hélice par rapport à l'autre, ce qui ouvre le canal. Des sous-
20 unités accessoires peuvent modifier la fonction du canal.

On a découvert l'utilité thérapeutique de matériels recombinants issus de l'ADN des nombreux canaux sodiques. Par exemple, US-A-5 132 296 (Cherksey) décrit des canaux sodiques purifiés qui se sont révélés utiles en tant
25 qu'outils diagnostiques et thérapeutiques.

Des isoformes de canaux sodiques sont divisées en "sous-familles". On utilise le terme "isoforme" pour désigner des protéines de canaux sodiques distinctes mais étroitement apparentées, c'est-à-dire celles ayant une
30 homologie d'acides aminés d'environ 60-80 %. Ces dernières manifestent également une forte homologie de fonctions. On utilise le terme "sous-familles" pour désigner des canaux sodiques distincts qui ont une homologie d'acides aminés d'environ 80-95 %. On utilise des combinaisons de plu-
35 sieurs facteurs pour déterminer les distinctions au sein

d'une sous-famille, par exemple, la rapidité d'un canal, la localisation chromosomique, les données d'expression, l'homologie avec d'autres canaux au sein d'une même espèce, et l'homologie avec un canal de la même sous-famille entre des espèces différentes. Une autre considération est une affinité pour la tétrodotoxine ("TTX"). La TTX est une toxine très puissante provenant du tétrodon, ou poisson-globe, qui bloque la transmission des influx nerveux le long des axones et dans les membranes excitables des fibres nerveuses. La TTX se fixe au canal sodique et arrête le flux des ions sodium.

Des études utilisant la TTX comme sonde ont éclairé beaucoup le mécanisme et la structure des canaux sodiques. Il existe trois sous-types de canaux sodiques qui sont définis par l'affinité pour la TTX, qui peut être mesurée par les valeurs de la CI_{50} : canaux sodiques sensibles à la TTX ($CI_{50} = 1-30$ nM), canaux sodiques insensibles à la TTX ($CI_{50} = 1-5$ μ M) et canaux sodiques résistants à la TTX ($CI_{50} \geq 50$ μ M).

Les potentiels d'action insensibles à la TTX ont été étudiés d'abord dans le muscle squelettique de rat (Redfern et coll., *Acta Physiol. Scand.* 82, 70-78 (1971)). Ces potentiels d'action ont été ensuite décrits dans d'autres tissus mammaliens, y compris le muscle squelettique mammalien de nouveau-né, le muscle cardiaque mammalien, des cellules de ganglions de la racine dorsale de souris in vitro et en culture, des cellules L6 et de muscle squelettique mammalien en culture [voir Rogart, *Ann. Rev. Physiol.* 43, 711-725 (1980)].

Les neurones des ganglions de la racine dorsale de rat possèdent à la fois des courants de canaux sodiques sensibles à la TTX ($CI_{50} \sim 0,3$ nM) et des courants de canaux sodiques résistants à la TTX ($CI_{50} \sim 100$ μ M), comme décrit par Roy et coll. dans *J. Neurosci.* 12, 2104-2111 (1992)]. On a également mesuré des courants sodiques

résistants à la TTX dans des ganglions pétreux et des ganglions plexiformes de rat (voir Ikeda et coll., *J. Neurophysiol.* 55, 527-539 (1986) et Stea et coll., *Neurosci.* 47, 727-736 (1992). Les électrophysiologistes pensent
5 qu'un autre canal sodique résistant à la TTX reste à découvrir.

Bien que l'on connaisse des ADNc provenant de cerveau, de coeur et de muscle squelettique de rat, l'identification et l'isolement d'ADNc à partir de tissu nerveux
10 sensitif périphérique, tel que des ganglions de la racine dorsale, ont été gênés par la difficulté du travail avec un tel tissu.

La présente invention fournit de nouvelles séquences purifiées et isolées d'acide nucléique, codant
15 pour des protéines de canaux sodiques de tissu nerveux, de préférence résistantes à la TTX, qui sont fortement exprimées dans des ganglions plexiformes et des ganglions de la racine dorsale de l'adulte, moins fortement exprimées dans le cerveau, la moelle épinière et les ganglions cervicaux
20 supérieurs, et ne sont pas exprimées dans le nerf sciatique, le coeur ou le muscle squelettique. Sous des formes actuellement préférées, les nouvelles séquences d'ADN comprennent des séquences d'ADNc codant pour la protéine de canal sodique de tissu nerveux de rat. Un aspect de la
25 présente invention est la sous-unité α de cette protéine de canal sodique.

L'invention a également pour objet l'ADN, l'ADNc et l'ARNm issus des séquences d'acide nucléique de l'invention, et l'ARNc issu de l'ARNm. En particulier, deux
30 séquences d'ADNc ensemble codent pour le canal sodique entier de tissu nerveux de rat.

La présente invention englobe également d'autres formes d'ADN, telles que de l'ADN génomique, de l'ADN préparé par synthèse chimique partielle ou totale à partir de
35 nucléotides, et de l'ADN comportant des délétions ou des

mutations.

Encore un autre aspect de l'invention est la nouvelle protéine de canal sodique de rat, résistante à la TTX, et des fragments de celle-ci, codés par l'ADN de la présente invention.

Un autre aspect de la présente invention consiste en des polynucléotides et oligonucléotides recombinants comprenant une séquence d'acide nucléique issue de la séquence d'ADN de la présente invention.

Un autre aspect de l'invention est un procédé de stabilisation de l'ADNc entier qui code pour la séquence de la protéine de l'invention.

Un autre aspect de l'invention comprend des vecteurs d'expression comprenant l'ADN de l'invention, des cellules hôtes transformées ou transfectées par ces vecteurs et une banque d'ADNc de ces cellules hôtes.

La présente invention englobe également un essai pour la détection d'inhibiteurs de la protéine de canal potassique, comprenant la mise en contact d'un composé, soupçonné être un inhibiteur, avec un canal sodique exprimé, et la mesure de l'activité du canal sodique.

On fournit en outre un procédé d'inhibition de l'activité du canal sodique résistant à la TTX, comprenant l'administration d'une quantité efficace d'un composé ayant une valeur CI_{50} de 10 μM ou moins.

De plus, on fournit des procédés d'utilisation de l'ADN pour la production d'anticorps monoclonaux et d'anticorps polyclonaux, destinés à être utilisés en tant que cibles moléculaires pour la découverte de médicaments, en tant que marqueurs hautement spécifiques d'antigènes spécifiques, comme molécules détectrices, dans des essais diagnostiques, et pour des utilisations thérapeutiques, telles que le soulagement de la douleur, en tant que sonde du canal PN5 dans un autre tissu mammalien, dans la conception d'agents thérapeutiques et dans le criblage de

thérapies.

La présente invention est illustrée à l'aide des dessins ci-annexés. Sur ces dessins:

Les figures 1A-E représentent la séquence native
5 d'ADNc de 5 908 nucléotides codant pour le canal sodique type 5 ("PN5") de rat (SEQ ID n° 1), issu de deux clones d'ADNc chevauchants, désignés par 26.2 et 1.18.

Les figures 2A-F représentent la séquence d'acides
10 déduite de PN5 (SEQ ID n° 2, représentée dans le code d'acides à trois lettres). Les figures 2G-H, représentant la séquence d'acides déduite de PN5 dans le code d'acides à une seule lettre, montrent également les domaines homologues (I-IV); les segments transmembranaires (S1-S6) supposés; l'acide conférant la
15 résistance à TTX (♦); les sites de N-glycosylation (●); le site de phosphorylation à la protéine kinase A (PKA) dépendant de CAMP (o) et le codon de terminaison (*).

La figure 3A représente une séquence de 856 paires
20 de bases pour le PN5 humain (SEQ ID n° 3). La figure 3B représente la comparaison de séquence d'acides du fragment hPN5 avec le PN5 de rat.

La figure 4 représente la séquence de la nouvelle sonde de domaine IV de canal sodique (SEQ ID n° 4).

Les figures 5A-E représentent la séquence de
25 5 334 nucléotides modifiée pour la stabilité et l'expression (SEQ ID n° 5). Les nucléotides 24 à 5 518 constituent la région de 5 295 pb codant pour une protéine à 1 765 acides.

La figure 6 représente la carte de clonage de PN5.

30 La présente invention concerne une séquence d'acide nucléique purifiée et isolée codant pour une nouvelle protéine de canal sodique mammalienne, de préférence résistante à la TTX. L'expression "ADN purifié et isolé" signifie de l'ADN qui est essentiellement exempt, c'est-à-dire
35 contient moins d'environ 30 %, de préférence moins d'environ

ron 10 %, et encore mieux de préférence moins d'environ 1 %, de l'ADN auquel est associé à l'état naturel l'ADN d'intérêt. Les techniques pour déterminer la pureté sont bien connues dans le domaine et comprennent, par exemple, 5 la cartographie de restriction, l'électrophorèse en gel d'agarose et la centrifugation en gradient de CsCl.

Le terme "ADN" est destiné à comprendre l'ADNc, ou ADN complémentaire, qui consiste en séquences d'ADN simple brin ou double brin, produites par transcription inverse 10 d'ARNm isolé à partir d'une cellule donneuse ou par synthèse chimique. Par exemple, le traitement d'ARNm par une transcriptase inverse, telle que la transcriptase inverse d'AMV ou la transcriptase inverse de M-MuLV, en présence d'une amorce oligonucléotidique, donnera un duplex ARN-ADN 15 pouvant être traité par de la RNase H, de l'ADN polymérase et de l'ADN ligase, pour donner de l'ADNc double brin. Si on le désire, l'ADNc double brin peut être dénaturé par des techniques classiques, telles qu'un chauffage, pour donner de l'ADNc simple brin. Le terme "ADNc" comprend de 20 l'ADN qui est une copie complémentaire de l'ARNm existant dans la nature, ainsi que des copies complémentaires de variants de l'ARNm existant dans la nature, qui ont la même activité biologique. Les variants comprennent, par exemple, des insertions, des délétions, des séquences à 25 codons dégénérés et des allèles.

Un "ARNc" correspondant à l'ARNm transcrit à partir d'une séquence d'ADN codant pour la sous-unité α d'une nouvelle protéine de canal sodique, de préférence résistante à la TTX, est envisagé par la présente invention. 30 Le terme "ARNc" désigne un ARN qui est une copie de l'ARNm transcrit par une cellule.

En particulier, l'invention englobe de l'ADN comportant les versions natives des séquences nucléotidiques indiquées sur les figures 1A-E (SEQ ID n° 1), désigné ici 35 par canal sodique type 5 (PN5). Les figures 1A-E repré-

sentent le produit de construction d'ADNc à 5 908 nucléotides, comprenant un cadre de lecture ouvert de 5 298 bases (y compris le codon d'arrêt) (SEQ ID n° 1). Le résidu nucléotidique 79 représente le site d'initiation et de traduction, et le résidu 5367 représente la fin du codon d'arrêt.

L'invention englobe également des versions modifiées de PN5, et en particulier la version représentée sur les figures 5A-E (SEQ ID n° 5). Le clone SalI-XbaI de 5 334 nucléotides est dépourvu de la plupart des séquences non traduites, du cadre de lecture ouvert de 5 298 nucléotides partant du nucléotide 24 et se terminant au nucléotide 5321. Les codons d'initiation et d'arrêt sont soulignés, comme le sont les mutations traductionnellement silencieuses au niveau des nucléotides 3932, 3935, 3941, 3944 et 3947, qui ont été introduites pour bloquer un réarrangement dans cette région pendant la croissance dans *E. coli*.

La séquence nucléotidique de SEQ ID n° 1 (figures 1A-E) correspond aux ADNc de rat. Une recherche d'homologie a montré que le canal sodique le plus étroitement apparenté se trouvait dans le canal cardiaque de rat, avec une homologie de 72,5 %. Les canaux les plus étroitement apparentés qui viennent ensuite sont rPN1, avec 72 %, et les types I et III de cerveau de rat, avec 71,8 % et 71,3 %, respectivement. Les homologies avec rPN3a, hPN3, rPN4, rPN4a, le type II de cerveau de rat et le muscle squelettique de rat vont chacune environ de 70 à 71 %.

En outre, un clone de 856 paires de bases (SEQ ID n° 3), tel que représenté sur la figure 3A, a été isolé à partir d'une banque d'ADNc de ganglions de la racine dorsale (GRD) humains, et est étroitement apparenté à la séquence d'acides aminés de PN5 de rat, avec une identité de 79 % et une homologie de 86 %. La séquence de PN5 humain couvre la région entre IIIS1 et l'interdomaine III/IV, qui

inclut la vanne d'inactivation rapide (à savoir IFM) qui est localisée dans l'interdomaine III/IV.

Le terme "banque d'ADNc" désigne une collection de clones, habituellement dans un bactériophage ou, moins
5 communément, dans des plasmides bactériens, contenant des copies d'ADNc de séquences d'ARNm issues d'une cellule donneuse ou d'un tissu donneur.

On pense que les homologues supplémentaires du nouveau canal sodique de rat, résistant à la TTX, décrit ici,
10 sont également exprimés dans un autre tissu mammalien.

L'analyse northern blot (exemple 5) indique que PN5 est codé par un transcrit de ~ 6,5 kb.

La séquence d'acides aminés déduite de PN5, représentée sur les figures 2A-F (SEQ ID n° 2), manifeste les
15 caractéristiques structurales primaires d'une sous-unité α d'un canal sodique dépendant du potentiel, résistant à la TTX. Les figures 2G-H représentent les domaines homologues (I-IV); les segments transmembranaires (S1-S6) supposés; l'acide aminé conférant la résistance à la TTX (\diamond); les
20 sites de N-glycosylation (\bullet); et les sites de phosphorylation à la PKA cAMP-dépendants (o). Des séquences d'ADN codant pour les polypeptides de protéine de canal sodique identiques, analogues ou variants allèles du système nerveux, par utilisation, au moins en partie, de codons dégénérés,
25 sont également envisagées par la présente invention.

Une caractéristique intéressante de cette séquence d'acides aminés déduite est que l'acide aminé qui est le plus responsable de la sensibilité à la TTX est localisé à la position 355 et n'est pas aromatique. Dans les canaux
30 sodiques de type du cerveau humain et de rat, dans le canal du muscle squelettique et dans PN1 et PN4, cet acide aminé est la tyrosine ou la phénylalanine, et ces canaux sont tous sensibles à la TTX. Dans PN3 et PN5, l'acide aminé est la sérine. Etant donné que PN3 est très
35 résistant à la TTX, l'implication est que PN5 est égale-

ment un canal résistant à la TTX. Le canal cardiaque comporte un résidu cystéine à cette position et est "insensible" à la TTX.

Bien que PN5 contienne toutes les caractéristiques distinctives d'un canal sodique dépendant du potentiel, il a des caractéristiques structurales spécifiques qui le distinguent des autres canaux sodiques. Par exemple, DIIS4 comporte 5 aminoacides basiques conservés dans tous les canaux sodiques qui pourraient jouer un rôle significatif dans les aspects de détection de potentiel de la fonction de canal. Dans PN5, le premier aminoacide basique est remplacé par un résidu alanine. De même, dans DIIIS4, PN5 comporte 5 aminoacides basiques au lieu des six qui sont présents dans d'autres séquences de canaux sodiques, le dernier résidu arginine étant remplacé par un résidu glutamine. Dans DIIIS3, le segment transmembranaire ne contient que 18 aminoacides, par opposition à 22 aminoacides dans les autres canaux. En outre, la courte boucle de liaison (4 aminoacides) entre S3 et S4 dans DIII est même plus courte, en raison d'une délétion de 3 aminoacides. Ce raccourcissement de S3 et de la boucle de liaison a été confirmé par la conception d'amorces dans la région appropriée de la séquence pour un essai de RT-PCR (réaction d'amplification en chaîne par polymérase-transcription inverse) à partir de GRD de rat, et séquençage du fragment d'ADN amplifié. Un tel essai a été effectué pour confirmer la séquence d'une autre région de PN5, dans la boucle DIVS5-S6, où il y avait une délétion d'un peptide de 8 aminoacides.

On a effectué une analyse de distribution tissulaire d'ARN, par une réaction d'amplification en chaîne par polymérase-transcription inverse (RT-PCR amorcée par oligonucléotide), à partir des systèmes nerveux périphérique et central de rat, en particulier à partir de GRD de rat. On a étudié huit types de tissus principaux quant à

l'expression des gènes PN5 spécifiques correspondants aux positions 5651-5903 de SEQ ID n° 1 (figures 1A-E). L'ARNm de PN5 était présent dans cinq des tissus étudiés: le cerveau, la moelle épinière, GRD, les ganglions plexiformes et les ganglions cervicaux supérieurs. PN5 n'était pas présent dans les tissus restants étudiés: le tissu de nerf sciatique, le tissu de muscle squelettique ou cardiaque. PN5 s'est révélé être le plus fort dans GRD et les ganglions plexiformes, ce qui a conduit les demandeurs à penser que le GRD était enrichi en PN5. PN5 présente des différences considérables d'abondance parmi un ensemble de tissus. PN5 a un gradient d'expression avec une forte expression dans les GRD. PN5 a un gradient d'expression comme les autres canaux, mais une distribution plus limitée.

L'invention comprend non seulement la protéine entière exprimée par les séquences d'ADNc de SEQ ID n° 1, 2 et 3, mais comprend également des fragments de protéines. Ces fragments peuvent être obtenus par coupure des protéines entières ou par l'utilisation de plus courtes séquences d'ADN ou de plus courts polynucléotides, pour exprimer le fragment recherché.

Tel qu'utilisé ici, le terme "polynucléotide" désigne une forme polymère de nucléotides de longueur quelconque, qu'il s'agisse de ribonucléotides ou de désoxyribonucléotides. Ce terme désigne seulement la structure primaire de la molécule. Ainsi, ce terme comprend de l'ADN simple brin et de l'ADN double brin, ainsi que de l'ARN double brin et de l'ARN simple brin. Il comprend également des formes modifiées, par exemple par méthylation et/ou par coiffage, et des formes non modifiées du polynucléotide.

En outre, le terme "polynucléotide" est destiné à comprendre un polynucléotide recombinant qui est d'origine génomique, d'ADNc, semi-synthétique ou synthétique, qui,

en vertu de son origine ou d'une manipulation, n'est pas associé à la totalité ou une partie du polynucléotide auquel il est associé dans la nature et/ou est lié à un polynucléotide autre que celui auquel il est lié dans la nature.

En conséquence, l'invention comprend également des polynucléotides qui peuvent être utilisés pour produire des polypeptides ayant une longueur d'environ 10 à 1 500, de préférence de 10 à 100 aminoacides. L'isolement et la purification de tels polypeptides recombinants peuvent être effectués par des techniques qui sont bien connues dans le domaine, par exemple des séparations chromatographiques préparatives ou la chromatographie d'affinité. En outre, des polypeptides peuvent également être produits par des moyens synthétiques qui sont bien connus dans la technique.

L'invention permet la manipulation de matériels génétiques par des techniques de recombinaison, pour produire des polypeptides qui possèdent les caractéristiques structurales et fonctionnelles de la nouvelle sous-unité α de canal sodique dépendant du potentiel, résistante à la TTX, rencontrée dans des nerf sensitifs. On peut utiliser la mutagénèse dirigée sur site pour obtenir de tels polypeptides recombinants. Par exemple, on peut insérer spécifiquement des oligonucléotides synthétiques, ou les mettre spécifiquement à la place d'oligonucléotides existants, dans le segment du gène d'intérêt, pour produire des gènes codant pour et exprimant un mutant spécifique. On peut également insérer des oligonucléotides dégénérés au hasard et on peut utiliser des techniques de visualisation de phages pour identifier et isoler des polypeptides ayant une propriété fonctionnelle d'intérêt.

En outre, la présente invention envisage des polynucléotides recombinants ayant une longueur d'environ 15 à 20 kb, de préférence de 10 à 15 kb, comprenant une

séquence d'acide nucléique issue de l'ADN de l'invention.

Le terme "issu" d'une séquence désignée se réfère à une séquence d'acide nucléique qui est composée d'une séquence d'environ au moins 6 à 8 nucléotides, de préférence d'au moins 10 à 12 nucléotides et en particulier d'au moins 15 à 20 nucléotides qui correspondent à une région de la séquence désignée, c'est-à-dire sont homologues ou complémentaires de celle-ci. La séquence dérivée n'est pas nécessairement dérivée physiquement de la séquence nucléotidique indiquée, mais peut être issue d'une autre manière, y compris, par exemple, la synthèse chimique ou la répllication d'ADN ou la transcription inverse, techniques qui sont basées sur l'information fournie par les séquences de base dans la ou les régions desquelles est issu le polynucléotide.

Un test d'expression néonatal a été effectué avec F11, une lignée de cellules de fusion conçue à partir de GRD de rat nouveau-né, soudée à une lignée cellulaire de souris, N18TG, provenant de Massachusetts General Hospital. F11 répond à des agents trophiques, tels que le NGF (facteur de croissance des neurones), par extension de dendrites. On a constaté que PN5 était présent à la fois dans F11 native et F11 traitée par NGF, ce qui a conduit les demandeurs à penser que le canal sodique était exprimé à l'état natif dans F11.

Une hybridation in situ d'ARNm de PN5 avec un tissu de GRD de rat donne une localisation principalement dans les petits neurones et les neurones moyens, avec une absence dans les neurones de grande taille.

On a également cartographié PN5 jusqu'à sa localisation cytogénétique sur des préparations de chromosomes de souris. PN5 est localisé sur le même chromosome que le canal cardiaque et PN3.

En général, les canaux sodiques comprennent une sous-unité α et deux sous-unités β . Les sous-unités β

peuvent moduler la fonction du canal. Toutefois, étant donné que la sous-unité α est tout ce qui est requis pour que le canal soit totalement fonctionnel, l'expression de l'ADNc dans SEQ ID n° 1 (figures 1A-E) donnera une protéine totalement fonctionnelle. Le gène codant pour la sous-unité β_1 dans le tissu de nerf périphérique s'est révélé être identique à celui rencontré dans le coeur, le cerveau et le muscle squelettique de rat. L'ADNc de la sous-unité β_1 n'est pas décrit ici, puisqu'il est bien connu dans la technique [voir Isom et coll., *Neuron* 12, 1183-1194 (1994)]. Toutefois, il doit être entendu qu'en combinant la séquence connue codant pour la sous-unité β_1 avec la séquence de la sous-unité α décrite ici, il est possible d'obtenir un canal sodique PNS complet, dépendant du potentiel, de préférence résistant à la TTX.

La présente invention comprend également des vecteurs d'expression comprenant l'ADN ou l'ADNc décrit plus haut, des cellules hôtes transformées par ces vecteurs d'expression, capables de produire le canal sodique de l'invention, ainsi que des banques d'ADNc comprenant de telles cellules hôtes.

Le terme "vecteur d'expression" signifie tout élément génétique, par exemple un plasmide, un chromosome, un virus, se comportant comme une unité autonome d'expression de polynucléotide dans une cellule ou étant rendu capable de répllication par insertion dans un chromosome d'une cellule hôte, et auquel est attaché un autre segment polynucléotidique, de manière à provoquer la répllication et/ou l'expression du segment attaché. Des vecteurs appropriés comprennent, mais sans se limiter à ceux-ci, des plasmides, des bactériophages et des cosmides. Les vecteurs contiendront les séquences polynucléotidiques qui sont nécessaires pour effectuer la ligature ou l'insertion du vecteur dans une cellule hôte désirée et pour effectuer l'expression du segment attaché. De telles séquences dif-

fèrent en fonction de l'organisme hôte ou comprendront des séquences de promoteurs pour effectuer la transcription, des séquences d'activateurs pour accroître la transcription, des séquences de sites de liaison au ribosome et
5 des séquences de terminaison de transcription et de traduction.

L'expression "cellule hôte" se réfère en général à des organismes procaryotes ou eucaryotes et comprend tout organisme apte à la transformation ou à la transfection,
10 qui est capable d'exprimer une protéine et peut être, ou a été, utilisé en tant que receveur de vecteurs d'expression ou d'autre ADN transféré. On peut également contraindre des cellules hôtes à exprimer une protéine par injection directe d'ARNc exogène, pouvant être traduit en la pro-
15 téine d'intérêt. Une cellule hôte préférée est l'ovocyte de *Xenopus*.

Le terme "transformé" se réfère à toute méthode connue pour l'insertion de séquences d'ADN ou d'ARN étranger dans une cellule hôte procaryote. Le terme "trans-
20 fecté" se réfère à toute méthode connue pour l'insertion de séquences d'ADN ou d'ARN étranger dans une cellule hôte eucaryote. De telles cellules transformées ou transfectées comprennent des cellules transformées ou transfectées de façon stable, dans lesquelles l'ADN inséré est rendu
25 capable de répllication dans la cellule hôte. Elles comprennent également des cellules exprimant temporairement, qui expriment pendant des durées limitées l'ADN ou l'ARN inséré. La technique de transformation ou de transfection dépend de la cellule hôte qui est transformée. Elle peut
30 comprendre l'encapsidation du polynucléotide dans un virus ainsi que l'introduction directe du polynucléotide, comme par exemple par lipofection ou micro-injection. La transformation ou la transfection peut entraîner l'incorporation de l'ADN inséré dans le génome de la cellule hôte ou
35 le maintien de l'ADN inséré dans la cellule hôte, sous

forme de plasmide. Des méthodes de transformation sont bien connues dans la technique et comprennent, mais sans se limiter à celles-ci, l'infection virale, l'électroporation, la lipofection et l'introduction directe provoquée par le phosphate de calcium.

Il doit être entendu que la présente invention est destinée à inclure d'autres formes de vecteurs d'expression, de cellules hôtes et de techniques de transformation servant à des fonctions équivalentes et bien connues dans la technique.

L'invention concerne également un essai d'inhibiteurs de la nouvelle protéine de canal sodique résistante à la TTX, comprenant la mise en contact d'un composé, soupçonné être un inhibiteur, avec un canal sodique exprimé, et la mesure de l'activité du canal sodique. Le composé peut être un composé pratiquement pur d'origine synthétique, mis en contact dans un milieu aqueux, ou le composé peut être une substance existant dans la nature, de sorte que le milieu d'essai est un extrait d'origine biologique, comme par exemple un extrait de cellules végétales, animales ou microbiennes. L'activité PN5 peut être mesurée par des méthodes telles que l'électrophysiologie (mise sous tension à l'aide de deux électrodes ou mise sous tension par patch clamp de cellules entières à l'aide d'une seule électrode), essais de flux d'ions guanidinium et essais de liaison à des toxines. Un "inhibiteur" est défini en général par la quantité entraînant une diminution de plus de 50 % de l'activité PN5, de préférence une diminution de plus de 70 % de l'activité PN5, encore mieux une diminution de plus de 90 % de l'activité PN5.

Il existe de nombreuses utilisations de l'invention, dont quelques unes sont décrites ci-dessous:

1. Sonde pour canaux mammaliens

Comme mentionné plus haut, on pense que des homologues supplémentaires du nouveau canal sodique de rat

résistant à la TTX, décrit ici, sont également exprimés dans un tissu mammalien, en particulier dans un tissu humain. On peut utiliser comme sonde les ADNc entiers de canaux sodiques PN5 de rat de la présente invention pour
5 découvrir s'il existe dans le tissu humain de nouveaux canaux sodiques supplémentaires, dépendants du potentiel, de préférence résistants à la TTX, et, s'ils existent, pour faciliter l'isolement des ADNc codant pour la protéine humaine.

- 10 Les homologues humains des canaux PN5 de rat, résistants à la TTX, peuvent être clonés à l'aide d'une banque d'ADNc de GRD humain. Les GRD humains sont obtenus lors d'une autopsie. On homogénéise le tissu congelé et on extrait l'ARN avec de l'isothiocyanate de guanidine
15 [Chirgwin et coll., *Biochemistry* 18, 5294-5299 (1979)]. L'ARN est fractionné en fonction de la taille sur un gradient de saccharose, pour l'enrichissement en ARNm de grande taille, car les sous-unités α du canal sodique sont codées par des transcrits de grande taille (7-11 kb). On
20 prépare de l'ADNc double brin en utilisant le nécessaire d'ADNc SuperScript Choice (GIBCO BRL) avec de l'oligo(dT) ou des amorces hexamères au hasard. On ligature les segments de liaison EcoRI sur l'ADNc double brin, qui est ensuite phosphorylé. On construit la banque d'ADNc par
25 ligature de l'ADNc double brin dans le vecteur bactériophage λ ZAP II (Stratagene), suivie d'encapsulation dans des particules phagiques.

- On étale les phages sur des plaques de 150 mm, sur un tapis de bactéries XLI-Blue MRF' (Stratagene) et on
30 fait des répliques des plages sur des membranes en Nylon Hybond N (Amersham). Les filtres sont hybridés avec des sondes d'ADNc de PN5 de rat, par des techniques classiques, et détectés par autoradiographie ou chimiluminescence. Le signal produit par les sondes PN5 de rat s'hy-
35 bridant avec des clones humains positifs, à une forte

stringence, devrait être plus fort que ceux obtenus avec des sondes de canal sodique de cerveau de rat s'hybridant avec ces clones. Les plages positives sont purifiées davantage par dilution limitante et criblées à nouveau par hybridation ou PCR. La cartographie de restriction et la réaction d'amplification en chaîne par polymérase identifieront des clones chevauchants qui peuvent être assemblés par des techniques classiques en l'homologue humain complet de PNS de rat. Le clone humain peut être exprimé par injection d'ARNc transcrit in vitro à partir du clone d'ADNc complet dans des ovocytes de *Xenopus*, ou par transfection d'une lignée cellulaire mammalienne avec un vecteur contenant l'ADNc lié à un promoteur approprié.

2. Anticorps contre PNS

Les polypeptides de l'invention sont très utiles pour la production d'anticorps dirigés contre PNS. Ces anticorps peuvent être utilisés dans une chromatographie d'affinité pour purifier des protéines recombinantes ou des polypeptides recombinants de canaux sodiques, ou ils peuvent être utilisés en tant qu'outil de recherche. Par exemple, des anticorps fixés à une molécule rapporteuse peuvent être utilisés dans des techniques de coloration histochimique, pour identifier d'autres tissus et d'autres types cellulaires dans lesquels les PNS sont présents, ou ils peuvent être utilisés pour identifier des régions fonctionnelles ou d'épitopes de la protéine de canal sodique de l'invention.

Les anticorps peuvent être monoclonaux ou polyclonaux et peuvent être produits par des techniques qui sont bien connues dans le domaine. Des anticorps polyclonaux sont produits comme suit: on utilise un conjugué immunogène, comprenant PNS ou un fragment de celui-ci, éventuellement lié à une protéine porteuse, pour immuniser un mammifère choisi, tel que la souris, le rat, la chèvre, etc. On recueille et traite selon des techniques connues

du sérum provenant du mammifère immunisé, pour séparer la fraction d'immunoglobuline.

Des anticorps monoclonaux sont produits par la technique classique d'hybridomes, basée sur celle rapportée par Kohler et Milstein dans *Nature* 256, 495-497 (1975). Des cellules spléniques sont obtenues à partir d'un animal hôte immunisé avec la protéine PN5 ou un fragment de celle-ci, éventuellement liées à une substance porteuse. Des cellules hybrides sont formées par fusion de ces cellules spléniques avec une lignée de myélome appropriée, et cultivées. Les anticorps produits par les cellules hybrides sont criblés en fonction de leur aptitude à se lier à des protéines PN5 exprimées.

On peut utiliser un certain nombre de techniques de criblage bien connues dans le domaine, comme par exemple des méthodes de criblage par essai immunoenzymatique direct ou indirect utilisant un corps adsorbé. On soumet ensuite les cellules hybrides produisant de tels anticorps à un reclonage ou à des conditions de forte dilution, afin de sélectionner une cellule hybride sécrétant une population homogène d'anticorps spécifiques de l'une ou l'autre protéine de PN5.

En outre, des anticorps peuvent être suscités par clonage et expression de séquences nucléotidiques ou de versions mutées de celles-ci, codant au moins pour les séquences d'acides aminés requises pour la fixation spécifique d'anticorps naturels, et ces protéines exprimées peuvent être utilisées en tant qu'immunogènes. Les anticorps peuvent comprendre l'immunoglobuline complète ou un fragment de celle-ci. Les anticorps peuvent être liés à un groupe rapporteur, comme décrit plus haut à propos des polynucléotides.

L'exemple 10 illustre la mise en pratique de la production d'un anticorps.

3. Cibles thérapeutiques pour des composés pour le traitement de troubles et essais de ceux-ci

La présente invention englobe également l'utilisation de la nouvelle sous-unité α de canal sodique dépendant du potentiel, de préférence résistante à la TTX, en tant que cible thérapeutique pour des composés destinés au traitement de troubles du système nerveux, sur la base des données de localisation de RT-PCR. Les troubles comprennent, mais sans se limiter à ceux-ci, l'épilepsie, une lésion due à une attaque cérébrale, une lésion cérébrale, la neuropathie diabétique, une lésion traumatique, la douleur neuropathique chronique et la neuropathie associée au SIDA.

4. Conception d'agents thérapeutiques à base de PN5 inhibiteur et essais de ceux-ci

La présente invention est également orientée vers l'inhibition de l'activité de PN5 dans des tissus du cerveau, de la moelle épinière, de ganglions de la racine dorsale, de ganglions plexiformes et de ganglions cervicaux supérieurs. Toutefois, il doit être entendu que des études ultérieures peuvent révéler que PN5 est présent dans d'autres tissus, et, en tant que tels, ces tissus peuvent également être des zones cibles. Par exemple, la détection d'ARNm de PN5 dans des ganglions plexiformes suggère que PN5 peut conduire des courants de sodium résistants à la TTX dans ces ganglions et d'autres ganglions sensitifs du système nerveux.

En outre, il s'est avéré que des protéines qui ne sont pas normalement exprimées dans certains tissus sont exprimées dans un état pathologique. En conséquence, la présente invention est destinée à englober l'inhibition de PN5 dans des tissus et des types cellulaires dans lesquels la protéine est normalement exprimée, et dans des tissus et des types cellulaires dans lesquels la protéine n'est exprimée qu'au cours d'un état pathologique.

Par exemple, on pense que des canaux sodiques résistants à la TTX jouent un rôle clé dans la transmission d'influx nerveux en relation avec des signaux sensoriels tels que la douleur et la pression. Cette information facilitera la conception d'agents thérapeutiques qui peuvent être ciblés vers une zone déterminée, telle qu'un tissu nerveux périphérique.

La protéine recombinante de la présente invention peut être utilisée pour l'essai d'agents thérapeutiques potentiels ayant l'aptitude à inhiber le canal sodique d'intérêt. En particulier, elle serait utile pour inhiber sélectivement la fonction de canaux sodiques dans des tissus nerveux périphériques responsables de la transmission de signaux de douleur et de pression sans affecter en même temps la fonction de canaux sodiques dans d'autres tissus, tels que le coeur et le muscle. Une telle sélectivité permettrait le traitement de la douleur sans provoquer d'effets secondaires dus à des complications cardiaques ou neuromusculaires. En conséquence, il serait utile d'avoir des séquences d'ADN codant pour des canaux sodiques exprimées sélectivement dans un tissu nerveux périphérique.

5. Analgésique

Les canaux sodiques dans le tissu nerveux périphérique jouent un rôle important dans la transmission d'influx nerveux, et sont par conséquent un outil pour comprendre la transmission de la douleur neuropathique. La douleur neuropathique se répartit en deux composantes: l'allodynie, dans laquelle un stimulus normalement non douloureux le devient, et l'hyperalgésie, dans laquelle un stimulus habituellement normalement douloureux devient extrêmement douloureux.

Dans des études de localisation tissulaire, l'ARNm de PN5 se trouve dans les petits neurones et les neurones moyens de GRD. L'ARNm de PN5 est également présent dans le cerveau et la moelle épinière. L'inhibition de ses activi-

tés peut faciliter la prévention de troubles tels que les maux de tête et les migraines. La possibilité d'inhiber l'activité de ces canaux sodiques, c'est-à-dire de réduire la conduction d'influx nerveux, affectera l'aptitude du
5 nerf à transmettre les influx douloureux. L'inhibition sélective des canaux sodiques dans des neurones sensitifs tels que GRD permettra l'arrêt d'influx douloureux sans effets secondaires indésirables provoqués par l'inhibition des canaux sodiques dans d'autres tissus tels que le cer-
10 veau et le coeur. En outre, certaines maladies sont provoquées par des canaux sodiques qui produisent des influx à une fréquence extrêmement élevée. La possibilité de réduire l'activité du canal peut alors éliminer ou atténuer la maladie. En conséquence, on peut cribler des com-
15 posés thérapeutiques potentiels par des méthodes bien connues dans la technique, pour découvrir s'ils peuvent inhiber l'activité du canal sodique recombinant de l'invention [Barram M. et coll., *Naun-Schmiedeberg's Archives of Pharmacology* 347, 125-132 (1993), et McNeal E.T. et coll.,
20 *J. Med. Chem.* 28, 381-388 (1985)]. Pour des études similaires avec le récepteur d'acétylcholine, voir Claudio et coll., *Science* 238, 1688-1694 (1987).

Par exemple, on peut soulager la douleur par inhibition de l'activité du nouveau canal sodique, de préfé-
25 rence résistant à la TTX, par administration d'une quantité thérapeutiquement efficace d'un composé ayant une CI_{50} d'environ 10 μM ou moins, de préférence $\leq 1 \mu M$. Des composés thérapeutiques potentiels sont identifiés sur la base de leur aptitude à inhiber l'activité de PN5. En con-
30 séquence, l'essai mentionné plus haut peut être utilisé pour identifier des composés ayant une CI_{50} thérapeutiquement efficace.

Le terme " CI_{50} " désigne la concentration d'un composé qui est requise pour inhiber à 50 % l'activité de PN5
35 exprimée, lorsque l'activité est mesurée par électro-

physiologie, essais de flux et essais de liaison à des toxines, comme mentionné plus haut.

6. Essais de diagnostic

Les techniques de biologie moléculaire de base utilisées dans la mise en oeuvre de la présente invention, telles que l'isolement d'ARN, d'ADN et de plasmides, la digestion par des enzymes de restriction, la construction et le sondage d'une banque d'ADNc, le séquençage de clones, la construction de vecteurs d'expression, la transformation de cellules, le maintien et le développement de cultures cellulaires et d'autres techniques générales, sont bien connues dans le domaine, et on peut trouver les descriptions de ces techniques dans des manuels généraux de laboratoire, tels que *Molecular Cloning: A Laboratory Manual* par Sambrook et coll. (Cold Spring Harbor Laboratory Press, 2^e édition, 1989).

Par exemple, les polynucléotides de l'invention peuvent être liés à une "molécule rapporteuse", pour la formation d'une sonde polynucléotidique utile pour des analyses northern et Southern blot et des hybridations in situ.

L'expression "molécule rapporteuse" désigne une entité chimique pouvant être détectée par un moyen de détection approprié, comprenant, mais sans se limiter à ceux-ci, des moyens spectrophotométriques, chimiluminescents, immunochimiques ou radiochimiques. Les polynucléotides de la présente invention peuvent être conjugués à une molécule rapporteuse par des techniques bien connues dans le domaine. Normalement, la molécule rapporteuse contient un groupe fonctionnel approprié à l'attachement au polynucléotide ou à l'incorporation dans celui-ci. Les groupes fonctionnels appropriés à l'attachement du groupe rapporteur sont habituellement des esters activés ou des agents d'alkylation. Des détails techniques pour l'attachement de groupes rapporteurs sont bien connus dans le

domaine; voir par exemple Matthews J.A., Batki A., Hynds C. et Kricka L.J., *Anal. Biochem.* 151, 205-209 (1985) et Engelhardt et coll., EP-A-0 302 175.

La présente invention est illustrée par les
5 exemples descriptifs et non limitatifs ci-après.

Abréviations

Les abréviations suivantes sont utilisées dans les exemples et ont les significations données ci-dessous.

ASB : albumine de sérum bovin

10 EDTA: acide éthylènediaminetétraacétique, sel tétra-sodique

GRD : ganglions de la racine dorsale

MEN : MOPS 20 mM, EDTA 1 mM, acétate de sodium 5 mM, pH 7,0

15 MOPS: acide 3-(N-morpholino)propanesulfonique (Sigma Chemical Company)

PN5 : canal sodique 5 de nerf périphérique

SDS : dodécylsulfate de sodium

SNP : système nerveux périphérique

20 Solution de Denhardt: 0,02 % d'ASB, 0,02 % de polyvinylpyrrolidone, 0,02 % de Ficoll (0,1 g d'ASB, 0,1 g de Ficoll, 0,1 g de polyvinylpyrrolidone pour 500 ml)

SSC : NaCl 150 mM, citrate de sodium 15 mM, pH 7,0

SSPE: NaCl 80 mM, phosphate de sodium 10 mM,

25 éthylènediaminetétraacétate 1 mM, pH 8,0

TEV : pince de tension à deux électrodes

TTX : tétrodotoxine (Sigma Chemical Company).

EXEMPLES

Les exemples ci-après illustrent la mise en pra-
30 tique de l'invention.

Matériels

Le plasmide pBK-CMV a été obtenu auprès de Stratagene (La Jolla, CA, USA); le plasmide pBSTA est décrit par Goldin et coll. dans *Methods in Enzymology*
35 (Rudy & Iverson éditeurs) 207, 279-297; le plasmide pCInéo

a été obtenu auprès de Promega (Madison, WI, USA); et le plasmide pCRII a été obtenu auprès d'Invitrogen (Carlsbad, CA, USA).

Le plasmide vecteur d'expression dans des ovocytes pBSTAcIIr a été construit à partir de pBSTA par insertion d'un lieu oligonucléotidique synthétique; le plasmide pKK232-8 a été obtenu auprès de Pharmacia Biotech (Piscataway, NJ, USA); le plasmide pCRII a été obtenu auprès d'Invitrogen (San Diego, CA, USA). Les lignées cellulaires de *E. coli* compétentes STBL2™ et SURE® ont été obtenues auprès de Gibco/BRL et Stratagene, respectivement.

EXEMPLE 1

Obtention d'ARN à partir de GRD, de cerveau et de moelle épinière de rat

Sous un microscope à dissection, on a prélevé les GRD lombaires n° 4 et n° 5 (L4 et L5), le cerveau et la moelle épinière de rats Sprague-Dawley mâles adultes anesthésiés. Les tissus ont été congelés dans de la glace carbonique et homogénéisés à l'aide d'un homogénéisateur Polytron; on a extrait l'ARN par la méthode à l'isothiocyanate de guanidine [voir Chomczynski et coll., *Anal. Biochemistry* 162, 156-159 (1987)]. On a dissous l'ARN total (5 µg de chaque échantillon) dans du tampon MEN contenant 50 % de formamide, 6,6 % de formaldéhyde et dénaturé à 65°C pendant 5-10 minutes. On a soumis l'ARN à une électrophorèse sur un gel d'agarose à 0,8 % contenant 8,3 % de formaldéhyde, dans du tampon MEN. Le tampon d'électrode était du tampon MEN contenant 3,7 % de formaldéhyde; on a fait fonctionner le gel à 50 V pendant 12-18 heures.

On a soumis à l'électrophorèse dans des colonnes parallèles du gel des marqueurs de taille moléculaire, comprenant les ARN ribosomiques 18S et 28S et des marqueurs d'ARN (GIBCO BRL). Leurs positions ont été détermi-

nées par coloration au bromure d'éthidium (0,5 µg/ml) de la colonne excisée, suivie d'une photographie sous lumière UV.

- Après l'électrophorèse, on a rincé le gel dans SSC 2x et on a transféré l'ARN sur une membrane de Duralose (Stratagene) avec SSC 20x, par action capillaire; la membrane a été chauffée sous vide pendant 1 heure à 80°C.

EXEMPLE 2

Sonde provenant de IIA de cerveau de rat

- On a synthétisé in vitro une sonde d'ARNc marquée au ^{32}P , complémentaire des nucléotides 4637-5868 de la séquence de sous-unité α de canal sodique IIA de cerveau de rat, avec de l'ARN polymérase de T7 (Pharmacia), en utilisant de l'ADN matrice de pEAF8 (Noda et coll., *Nature* 320, 188-192 (1986)) qui avait été linéarisé à l'aide de *Bst*EII.

- On peut trouver des protocoles pour chaque technique mentionnée ci-dessus dans *Molecular Cloning: A Laboratory Manual* par Sambrook et coll. (Cold Spring Harbor Laboratory Press, 2^e édition, 1989).

EXEMPLE 3

Hybridation d'ARN avec la sonde provenant de IIA de cerveau de rat

- La membrane de l'exemple 1 a été préhybridée pendant 16 heures à 42°C dans 50 % de formamide, SSC 5x, phosphate de sodium 50 mM, pH 7,1, solution de Denhardt 1x, 0,5 % de SDS et 1 mg/ml d'ADN de spermatozoïdes de saumon cisailé, dénaturé par la chaleur. La membrane a été hybridée pendant 18 heures à 42°C dans 50 % de formamide, SSC 5x, phosphate de sodium 50 mM, pH 7,1, solution de Denhardt 1x, 0,5 % de SDS et 200 µg/ml d'ADN de spermatozoïdes de saumon cisailé, dénaturé par la chaleur, avec la sonde d'ARNc marquée au ^{32}P (environ $1-3 \times 10^6$ cpm/ml) décrite dans l'exemple 2.

On a rincé la membrane pendant 20 minutes avec SSC 2x, 0,1 % de SDS, à la température ambiante, et on l'a ensuite lavée successivement avec SSC 2x, 0,1 % de SDS à 55°C pendant 30 minutes; SSC 0,2x, 0,1 % de SDS à 65°C pendant 30 minutes; SSC 0,2x, 0,1 % de SDS à 70°C pendant 30 minutes et SSC 0,2x, 0,1 % de SDS, 0,1 % de pyrophosphate de sodium à 70°C pendant 20 minutes. Le filtre a été exposé contre un film Kodak X-omat AR à -80°C, avec des écrans amplificateurs, pendant jusqu'à 2 semaines.

La sonde de pEAF8 s'est hybridée avec les ARNm dans l'échantillon de GRD, avec des tailles de 11 kb, 9,5 kb, 7,3 kb et 6,5 kb, estimées sur la base de leurs positions par rapport aux étalons.

EXEMPLE 4

15 Nouvelle sonde de domaine IV de canal sodique

La sonde a été obtenue comme suit: on a effectué une RT-PCR sur de l'ARN isolé à partir de GRD de rat, en utilisant des amorces oligonucléotides dégénérées qui étaient conçues sur la base des homologies entre les canaux sodiques connus dans le domaine IV. Les produits de domaine IV ont été clonés dans un vecteur plasmidique, transformés dans *E. coli*, et on a isolé des colonies individuelles. Les produits de PCR spécifiques du domaine IV, obtenus à partir de plusieurs de ces colonies, ont été séquencés individuellement. La nouvelle séquence de domaine IV clonée était la suivante (SEQ ID n° 4):

```

1   CTCAACATGG TTACGATGAT GGTGGAGACC GACGAGCAGG GCGAGGAGAA
51  GACGAAGGTT CTGGGCAGAA TCAACCAGTT CTTTGTGGCC GTCTTCACGG
101 GCGAGTGTGT GATGAAGATG TTCGCCCTGC GACAGTACTA TTTCACCAAC
30 151 GGCTGGAACG TGTTGCACTT CATAGTGGTG ATCCTGTCCA TTGGGAGTCT
201 GCTGTTTCT  GCAATCCTTA AGTCACTGGA AAACACTTTC TCCCCGACGC
251 TCTTCCGGGT CATCCGTCTG GCCAGGATCG GCCGCATCCT CAGGCTGATC
301 CGAGCAGCCA AGGGGATTCG CACGCTGCTC TTCGCCCTCA TGATGTCCCT
351 GCCCGCCCTC TTCAACATCG GCCTCCTCCT CTTCTCGTC ATGTTTCATCT
35 401 ACTCCATCTT CGGCATGGCC AGCTTCGCTA ACGTCGTGGA CGAGGCCGGC

```

451 ATCGACGACA TGTTCAACTT CAAGACCTTT GGCAACAGCA TGCTGTGCCT
 501 GTTCCAGATC ACCACCTCGG CCGGCTGGGA CGGCCTCCTC AGCCCCATCC
 551 TCAACACGGG GCCTCCCTAC TGCGACCCCA ACCTGCCCAA CAGCAACGGC
 601 TCCCGGGGGA ACTGCGGGAG CCCGGCGGTG GGCATCATCT TCTTCACCAC
 5 651 CTACATCATC ATCTCCTTCC TCATCGTGGT CAACATGTAT ATCGCAGTCA
 701 TC

Cette séquence a été marquée au ^{32}P par amorçage au hasard.

EXEMPLE 5

10 Hybridation d'ARN avec la nouvelle sonde 3'-UTR de canal sodique

On a préparé une empreinte northern blot avec 10 μg d'ARN total provenant de cerveau, de moelle épinière et de GRD de rats. L'empreinte a été hybridée avec une sonde
 15 d'ARNc provenant de la région non traduite en 3' (3'-UTR). La 3'-UTR a été clonée dans le vecteur pSP73, et l'ARNc a été transcrit à l'aide d'un nécessaire Trans Probe T (Pharmacia Biotech) et de [^{32}P]-UTP. L'empreinte a été préhybridée pendant 2 heures à 65°C dans une solution con-
 20 tenant SSC 5x, solution de Denhardt 1x, 0,5 % de SDS, phosphate de sodium 50 mM, pH 7,1, 1 mg/ml d'ADN de spermatozoïdes de saumon et 50 % de formamide. L'hybridation a été effectuée à 45°C pendant 18 heures dans la solution ci-dessus, mis à part que l'ADN de spermatozoïdes de sau-
 25 mon a été inclus à une concentration de 200 $\mu\text{g}/\text{ml}$ et qu'on a ajouté la sonde marquée au ^{32}P à raison de $7,5 \times 10^5$ cpm/ml de solution. L'empreinte a été ensuite lavée trois fois avec SSC 2x et 0,1 % de SDS à la température ambiante, une fois avec SSC 0,2x et 0,1 % de SDS à
 30 65°C pendant 20 minutes, et une fois avec SSC 0,2x, 0,1 % de SDS et 0,1 % de pyrophosphate de sodium à 65°C pendant 20 minutes. L'empreinte a été analysée sur un appareil PhosphoImager (BioRad), après une exposition de 2 jours. Les résultats ont indiqué qu'il y avait un signal de bande
 35 de ~ 6,5 kb, présent dans le cerveau, seulement dans la

colonne contenant de l'ARN provenant de GRD. En raison de la faible abondance d'ARN de PN5, comme l'a montré l'essai de RT-PCR, la bande de 6,5 kb n'était pas décelable dans le cerveau ni la moelle épinière.

5

EXEMPLE 6

Construction et criblage d'une banque d'ADNc provenant de GRD de rat

On a préparé une banque d'ADNc adaptée à *EcoRI* à partir d'ARN poly(A)⁺ de GRD de rat Sprague-Dawley mâle adulte normal, en utilisant le système SuperScript Choice (GIBCO BRL). L'ADNc (> 4 kb) a été sélectionné par fractionnement en gradient de saccharose, comme décrit par Kieffer, *Gene* 109, 115-119 (1991). L'ADNc a été ensuite ligaturé dans le vecteur Zap Express (Stratagene) et encapsidé à l'aide de l'extrait d'encapsidation dans λ Gigapack II XL (Stratagene). De même, on a synthétisé une banque d'ADNc de GRD de > 2 kb.

Les phages ($3,5 \times 10^5$) ont été criblés par hybridation sur filtre avec une sonde marquée au ^{32}P [rBIIa, bases 4637-5868 comme suit, d'Auld et coll., *Neuron* 1, 449-461 (1988)]. Les filtres ont été hybridés dans 50 % de formamide, SSPE5x, solution de Denhardt 5x, 0,5 % de SDS, 250 $\mu\text{g/ml}$ d'ADN de spermatozoïdes de saumon cisailé, dénaturé, et phosphate de sodium 50 mM, à 42°C, et lavés dans SSC 0,5x/0,1 % de SDS, à 50°C.

Des empreintes Southern blot de plasmides digérés par *EcoRI* ont été hybridées avec la sonde d'ADN marquée au ^{32}P (SEQ ID n° 4). Les filtres ont été ensuite hybridés à 42°C dans 50 % de formamide, SSC 6x, solution de Denhardt 5x, 0,5 % de SDS et 100 $\mu\text{g/ml}$ d'ADN de spermatozoïdes de saumon cisailé, dénaturé, et ont été lavés dans SSC 0,1x/0,1 % de SDS, à 65°C.

On a excisé des clones positifs in vivo dans pBK-CMV, en utilisant le système ExAssist/XL0LR (Stratagene).

EXEMPLE 7

Clones et analyse des nucléotides

On a isolé des clones d'ADNc, 26.2 et 25.1, à partir de la banque d'ADNc, et on a isolé le clone 1.18 à partir de la banque d'ADNc de GRD de > 2 kb. Par analyse de la séquence, 26.2 est apparu comme étant un ADNc complet codant pour un nouveau canal sodique, et 25.1 s'étendait du domaine II à la 3'-UTR. Toutefois, chacun comportait une délétion qui tronquait la région codante. Le clone 1.18 comportait la région non traduite en 3', en plus de l'extrémité C-terminale de la séquence d'acides déduite de PN5. Le produit de construction dans le vecteur d'expression pBSTACIIr consistait en séquences provenant de 26.2 et 1.18.

On a obtenu l'homologie de PN5 avec les autres canaux sodiques connus en utilisant le programme GAP/Best Fit (GCG) :

	Canal	% de similarité	% d'identité
	PN3a	71	54
20	hPN3	71	55
	PN4	71	53
	PN4a	71	53
	PN1	72	55
	Type I de cerveau de rat	72	55
25	Type II de cerveau de rat	71	54
	Type III de cerveau de rat	71	54
	Canal cardiaque de rat	73	56
	Canal de muscle squelettique de rat	71	53

30 Stabilisation de l'ADN complet de PN5

A. Milieux, lignées cellulaires de *E. coli* et conditions de culture

La culture de fragments de PN5 a pu être réalisée dans des conditions classiques, mais la culture de plasmides contenant des produits de construction complets de

PN5 (dans pCInéo, pBSTAcIIr et d'autres vecteurs) ne pouvait pas être effectuée sans utilisation de milieux de culture spéciaux, de conditions particulières et de souches spéciales de *E. coli*. Les conditions suivantes se sont révélées être optimales: (1) utilisation de *E. coli* STBL2™ pour la transformation primaire après des réactions de ligature et pour une culture à grande échelle; (2) le milieu solide était FM 0,5x (voir ci-dessous) + LB 1x (1 % de tryptone, 0,5 % d'extrait de levure, 0,5 % de NaCl) + 15 g/l de gélose, ou FM 1x + LB 0,5x; (3) le milieu liquide était de façon optimale FM 1x + LB 0,5x; (4) on a utilisé 100 µg/ml de carbénicilline pour tous les milieux, étant donné qu'elle était moins rapidement métabolisée que l'ampicilline; (5) la température pour la croissance ne devait pas être supérieure à 30°C, normalement de 24-26°C; cela demandait de plus longues durées de culture que celles normalement utilisées, allant de 24 à 72 heures.

Milieu de congélation 2x (FM 2x):

20	K_2HPO_4	12,6 g
	Citrates trisodique	0,9 g
	$MgSO_4 \cdot 7H_2O$	0,18 g
	$(NH_4)_2SO_4$	1,8 g
	KH_2PO_4	3,6 g
	Glycérol	88 g
25	H_2O complément à 1 l	

FM 2x et les composants restants du milieu sont préparés séparément, stérilisés par passage à l'autoclave, refroidis jusqu'à au moins 60°C, et ajoutés ensemble pour la formation du milieu final. La carbénicilline est préparée à raison de 25 mg/ml de H_2O et stérilisée par filtration. FM 2x a été décrit en premier lieu pour la préparation de suspensions de base congelées de cellules bactériennes [Practical Methods in Molecular Biology, Schleif R.F. et Wensing P.C., Springer-Verlag, New York (1981), p. 201-202].

B. Vecteurs d'expression

Afin d'obtenir une stabilité accrue de l'ADNc complet, on a modifié le vecteur d'expression pBSTAcIIr d'ovocytes pour réduire le nombre de copies de plasmides lors de la culture dans *E. coli* et pour réduire la poursuite de la transcription à partir de séquences du vecteur, qui pourrait conduire à une expression cryptique toxique de la protéine PN5 [Brosius J., *Gene* 27, 151-160 (1984)]. On a fait digérer pBSTAcIIr par *PvuII*. Le fragment de 755 pb, contenant le promoteur de T7, la 5'-UTR de β -globine, le site de clonage multiple, la 3'-UTR de β -globine et le promoteur de T3, a été ligaturé avec le fragment de 3,6 kb contenant l'origine de réplication, le gène de résistance à l'ampicilline, les terminateurs de transcription $rrnBT_1$ et $rrnBT_1T_2$ de pKK232-8, qui avait été digéré complètement par *SmaI* et digéré partiellement par *PvuII* et traité par de la phosphatase intestinale de crevette pour empêcher l'autoligature. Le plasmide résultant, dans lequel l'orientation du fragment pBSTA était telle que le promoteur de T7 était le plus rapproché du terminateur $rrnBT_1$, a été identifié par cartographie de restriction, est dénommé pHQ8. Comme dans le cas de pBSTA, le sens de la transcription du gène de résistance à l'ampicilline et de l'origine de réplication de pHQ8 est opposé à celui de la cassette d'expression du gène, et la présence du terminateur $rrnBT_1$ devrait réduire toute poursuite de lecture restante à partir du vecteur, dans la cassette d'expression conduite par le promoteur de T7.

C. Assemblage d'ADNc complet pour l'expression

Etant donné que pBK-CMV.26.2 comporte une délétion de 58 pb (correspondant aux pb 4346 à 4403 de SEQ ID n° 1) et que la séquence de pBK-CMV.1.18 commence au niveau de la pb 4180 de SEQ ID n° 1, pBK-CMV.1.18 pouvait être utilisé pour "réparer" pBK-CMV.26.2. On a mis au point une stratégie pour assembler un ADNc complet provenant des

clones pBK-CMV.26.2 et pBK-CMV.1.18 en trois segments, en tronquant la 5'-UTR et la 3'-UTR et en introduisant des sites de restriction uniques aux extrémités 5' et 3' dans le processus. L'extrémité 5' a été engendrée par PCR à partir de 26.2, en tronquant la 5'-UTR, par incorporation d'un site *SalI* juste en amont du codon d'initiation. Le segment central était un fragment de restriction provenant de 26.2. L'extrémité 3' a été préparée par PCR de chevauchement à partir de 26.2 et 1.18, et incorporation d'un site *XbaI* juste en aval du codon d'arrêt. On a fait digérer ces segments au niveau de sites uniques de restriction, et on les a assemblés dans pBSTAcIIr. Ce produit de construction semblait avoir une séquence correcte, mais, après reclonage en tant que fragment *SalI*-*XbaI* dans pCInéo, on a trouvé deux types d'isolats, un comportant une délétion et un comportant une insertion de 8 pb. Un nouvel examen du clone pBSTAcIIr a montré que la séquence était "mixte" dans cette région, de sorte que le clone devait avoir été réarrangé. On a découvert que l'insertion de 8 pb était une répétition de l'un des membres d'une duplication de 8 pb dans la séquence native, formant une triple répétition de 8 pb dans l'isolat réarrangé. De nombreuses tentatives de clonage ont inévitablement donné naissance à ce réarrangement. On a utilisé la PCR de chevauchement pour introduire des mutations silencieuses dans l'une des répétitions de 8 pb, et un fragment contenant cette région a été inclus lorsque la région codant pour PN5 a été assemblée dans HQ8, la version à faible nombre de copies de pBSTAcIIr, pour donner le plasmide HR-1. Cette séquence s'est avérée être stable (voir figures 5A-E, SEQ ID n° 5).

On a préparé le fragment de l'extrémité 5' par PCR, en utilisant comme matrice l'ADN de pBK-CMV.26.2 et les amorces 4999 (CTTGGTCTGACTCTAGATCAGGGTGAAGATGGAGGAG; site *SalI* souligné, homologie avec PN5 en italiques, correspon-

dant aux pb 58-77 de SEQ ID n° 1; codon d'initiation en gras) et 4927 (GTCTTCTAGATGAGGGTTCAGTCATTGTG, correspondant aux pb 1067 à 1047 de SEQ ID N°1), suivie de purification en gel, digestion par *SalI* et *KpnI* (site *KpnI* au niveau des pb 1003-1008, SEQ ID n° 1) et purification en gel.

Le fragment central de 3,1 kb a été préparé par digestion de l'ADN de pBK-CMV.26.2 par *KpnI* et *AatII* (site *AatII* à 4133-4138), suivie d'une purification en gel.

Le fragment de l'extrémité 3' a été préparé comme suit: la PCR à l'aide des amorces 4837 (TCTGGGAAGTTTGAAG, correspondant aux pb 3613 à 3629 de SEQ ID n° 1) et 4931 (GACCACGAAGGCTATGTTGAGG, correspondant aux pb 4239 à 4218 de SEQ ID n° 1) sur l'ADN de pBK-CMV.26.2 en tant que matrice a donné un fragment de 0,6 kb. La PCR, à l'aide des amorces 4930 (CCTCAACATAGCCTTCGTGGTC, correspondant aux pb 4218 à 4239 de SEQ ID n° 1) et 4929

(GTCTTCTAGATGAGGGTTCAGTCATTGTG, site *XbaI* souligné, homologie avec PNR en italiques, correspondant aux pb 5386 à 5365 de SEQ ID n° 1, codon d'arrêt en gras) sur l'ADN de pBK-CMV.1.18 en tant que matrice, a donné un fragment de 1,2 kb, introduisant un site *XbaI* à 7 pb du codon d'arrêt. L'extrémité 3' du fragment 4837-4931 complète donc exactement l'extrémité 5' du fragment 4930-4929. Ces deux fragments ont été purifiés en gel, et une fraction de chaque a été réunie en tant que matrice dans une réaction PCR, à l'aide des amorces 4928 (CAAGCCTTTGTGTTTCGAC, correspondant aux pb 4084 à 4101 de SEQ ID n° 1) et 4929, pour donner un fragment de 1,3 kb. Ce fragment a été purifié en gel, digéré par *AatII* et *XbaI*, et le fragment de 1,2 kb a été purifié en gel.

Le fragment de l'extrémité 3' a été cloné dans pBSTAcIIr digéré par *AatII* et *XbaI*. Un isolat a été digéré par *SalI* et *KpnI* et ligaturé avec le fragment de l'extrémité 5'. Le plasmide résultant, après vérification de la séquence, a été digéré par *KpnI* et *AatII* et ligaturé avec

le fragment central de 3,1 kb, pour donner pBSTAcIIr.PN5 (clone 21). pBSTAcIIr.PN5 (clone 21) a été digéré par *Sall* et *XbaI*, pour libérer le fragment PN5 de 5,3 kb, qui a été cloné dans pCInéoII digéré par *Sall* et *XbaI*. On a trouvé
 5 de multiples isolats, dont GPII-1, qui a été séquencé complètement, était caractéristique et contenait un insert de 8 pb. Cet insert CAGAAGAA, après la pb 3994 de SEQ ID n° 1, convertissait la répétition directe de cette
 10 séquence en cet emplacement en une triple répétition directe, provoquant un déplacement dans le cadre de lecture. Dans une tentative pour réparer ce défaut, on a fait digérer pBSTAcIIr.PN5 (clone 21) par *NheI* (pb 2538-2543, SEQ ID n° 1) et *XhoI* (pb 4828-4833, SEQ ID n° 1), pour
 15 obtenir un fragment de 6,2 kb, et avec *AatII* et *XhoI*, pour obtenir un fragment de 0,7 kb qui a été ligaturé avec le fragment de 1,6 kb résultant de la digestion de pBK-CMV.26.2 par *AatII* et *NheI*. Bien qu'on n'ait trouvé aucun isolat totalement correct, un isolat, HA-4, comportait seulement un seul changement de base, une délétion de
 20 la C à la pb 4827 (SEQ ID n° 1), adjacente au site *XhoI*.

Afin d'éviter que se produise le réarrangement par insertion de 8 pb, on a introduit trois mutations silencieuses dans la répétition en 5', et deux mutations supplémentaires dans une série de T seraient également intro-
 25 duites, comme indiqué ci-dessous (pb 3982 à 4014, SEQ ID n° 1; sites de mutation soulignés, répétitions de 8 pb dans la séquence native en italiques):

séquence native GAC ATT TTT ATG ACA GAA GAA CAG AAG AAA TAT
 Asp Ile Phe Met Thr Glu Glu Gln Lys Lys Tyr
 30 séquence mutante GAC ATC TTC ATG ACT GAG GAG CAG AAG AAA TAT

Etant donné que l'isolat HA-4 comportait la séquence répétée directe native [par opposition à, par exemple, pBSTAcIIr.PN5 (clone 21)] et que la région voisine du défaut du site *XhoI* n'était pas impliquée, on l'a
 35 utilisé comme ADN matrice pour les réactions PCR sui-

vantes. L'amorce P5-3716S

(CCGAAGCCAATGTAACATTAGTAATTACTCGTG, correspondant aux pb 3684 à 3716, SEQ ID n° 1) a été appariée avec l'amorce P5-3969AS (GCTCCTCAGTCATGAAGATGTCTTGGCCACCTAAC, correspondant aux pb 4003 à 3969, SEQ ID n° 1, les bases mutées sont soulignées), pour donner un produit de 320 pb.

L'amorce P5-4017S

(GGCCAAGACATCTTCATGACTGAGGAGCAGAAGAAATATTAC, correspondant aux pb 3976 à 4017, SEQ ID n° 1; les bases mutées sont

soulignées) a été appariée avec l'amorce P5-4247AS

(CTCAAAGCAAAGACTTTGATGAGACACTCTATGG, correspondant aux pb 4280 à 4247, SEQ ID n° 1), pour donner un produit de 305 pb. L'extrémité 3' du fragment de 320 pb avait donc un segment de 28 pb correspondant exactement à l'extrémité 5'

du fragment de 305 pb. Les deux bandes ont été purifiées en gel, et une fraction de chaque a été réunie, dans une nouvelle réaction PCR, avec les amorces P5-3716S et

P5-4247AS, pour donner un produit de 597 pb, qui a été cloné T/A dans le vecteur pCRII. On a trouvé l'isolat

HO-7, qui avait la séquence recherchée. On a effectué une quadruple ligation pour assembler le PN5 modifié complet:

On a fait digérer par *SalI* et *XbaI* le vecteur d'expression HQ-8 dans des ovocytes, pour obtenir un fragment vecteur de 4,4 kb, GPII-1; on a fait digérer GPII-1 par

SalI et *MluI*, pour obtenir un fragment de 3,8 kb contenant la moitié 5' de PN5; on a fait digérer HO-7 par *MluI* (pb 3866 à 3871, SEQ ID n° 1) et *AatII*, pour obtenir un fragment de 0,3 kb contenant la région répétée de 8 pb mutante de PN5; on a fait digérer GPII-1 par *AatII* et *XbaI*, pour

obtenir le segment 3' restant de 1,3 kb de PN5. Une partie du mélange réactionnel de ligation a été transformé dans des cellules Stable 2 de *E. coli*. Parmi les isolats de 9,6 kb contenant tous les quatre fragments, HR-1 a été séquencé et s'est révélé comporter la séquence de 5,4 kb recherchée. Ces isolats se développaient bien et ne mani-

festaient pas de tendance au réarrangement. La séquence de cette version construite de PN5 est représentée sur les figures 5A-E (SEQ ID n° 5).

EXEMPLE 8

5 PNR humain

On a isolé un clone de 856 pb (figure 3A, SEQ ID n° 3) à partir d'une banque d'ADNc de ganglions de la racine dorsale (GRD) humains, qui est très étroitement apparenté à PN5 de rat, avec une identité de 79 % en ce qui concerne la séquence d'acides aminés. La séquence de PN5 humain couvre la région entre IIIS1 et l'interdomaine III/IV, qui comprend la vanne d'inactivation rapide (à savoir IFM), qui est localisée dans l'interdomaine III/IV.

On a construit la banque d'ADNc de GRD à partir d'ARN total de GRD 4 et 5 lombaires qui était amorcé au hasard. L'ADNc du premier brin a été synthétisé à l'aide de transcriptase inverse SuperScript II (GIBCO BRL), et la synthèse du second brin a été effectuée à l'aide d'ADN polymérase de T4. Des segments de raccord EcoRI ont été ligaturés avec les extrémités des ADNc double brin, et les fragments ont été clonés dans le vecteur ZAP II (Stratagene). La banque a été criblée à l'aide de PN3 de rat marqué à la digoxigénine, PN1 de rat et des sondes hH1 de coeur humain. Les clones positifs ont été séquencés et comparés à des séquences connues de canaux sodiques humains et de rat. Seul le clone précité a été identifié en tant que séquence de PN5 humain.

	Canal	% de similarité	% d'identité
	Cerveau humain (HBA)	76	69
30	Coeur humain (hH1)	81	74
	Coeur atypique humain	60	52
	Muscle squelettique humain	80	71
	Neuroendocrine humaine	78	71
	PN3 humain	77	70
35	PN1 de rat	79	72

Canal	% de similarité	% d'identité
PN3 de rat	78	71
PN4 de rat	78	70
PN5 de rat	86	79

- 5 La figure 3B compare la séquence d'acides aminés du fragment hPN5 avec la séquence d'acides aminés de PN5 de rat, dans la région appropriée.

EXEMPLE 9

Distribution tissulaire par RT-PCR

- 10 On a isolé à partir de rats Sprague-Dawley mâles adultes normaux, anesthésiés, des tissus de cerveau, de moelle épinière, de GRD, de ganglions plexiformes, de ganglions cervicaux supérieurs, de nerf sciatique, de cœur et de muscle squelettique, et on les a conservés à -80°C.
- 15 On a isolé l'ARN de chaque tissu à l'aide de RNazol (Tel-Test Inc.). L'ADNc amorcé au hasard a été transcrit par transcription inverse à partir de 500 mg d'ARN provenant de chaque tissu. L'amorce directe (CAGATTGTGTTCTCAGTACATTCC) et l'amorce inverse (CCAGGTGTCTAACGAATAAATAGG) ont été conçues à partir de la
- 20 région non traduite en 3', pour donner un fragment de 252 paires de bases. Les paramètres de cycle étaient 94°C/2 minutes (dénaturation), 94°C/30 s, 65°C/30 s et 72°C/1 min (35 cycles) et 72°C/4 min. Les produits de
- 25 réaction ont été analysés sur un gel d'agarose à 4 %.
- Un témoin positif et un témoin sans matrice ont été également inclus. On a également amplifié par PCR de l'ADNc provenant de chaque tissu, en utilisant des amorces spécifiques de la glyceraldéhyde-3-phosphate déshydrogénase, pour montrer la viabilité de la matrice, comme décrit
- 30 par Tso et coll., *Nucleic. Acid Res.* 13, 2485-2502 (1985).
- Le profil de distribution tissulaire de rPN5 par analyse par RT-PCR d'ARN provenant de tissus choisis de rat, était le suivant:

	<u>Tissu</u>	<u>RT-PCR (35 cycles)</u>
	Cerveau	+
	Moelle épinière	+
	GRD	+++
5	Ganglions plexiformes	+++
	Ganglions cervicaux supérieurs	+
	Nerf sciatique	-
	Coeur	-
	Muscle squelettique	-
10	F11 non traité	+
	F11 traité	+

PN5 a été également détecté après seulement 25 cycles (24 + 1) dans les mêmes cinq tissus que ci-dessus, en la même abondance relative.

15 EXEMPLE 10

Anticorps

Un peptide synthétique (26 aminoacides dans les interdomaines II et III - résidus 977 à 1002) a été conjugué avec KLH et l'anticorps a été suscité chez des lapins. L'antisérum a été ensuite purifié par affinité. PN5 constitue une sous-famille de nouveaux gènes de canaux sodiques; ces gènes sont différents de ceux détectables avec d'autres sondes (par exemple les sondes PEA8 et PN3).

25 Il doit être bien entendu que la description qui précède n'a été donnée qu'à titre illustratif et non limitatif et que toutes variantes ou modifications peuvent y être apportées sans sortir pour autant du cadre général de présente invention, tel que défini dans les revendications
30 ci-annexées.

LISTE DES SEQUENCES

(1) INFORMATIONS GENERALES:

(I) DEMANDERESSE:

- 5 (A) NOM: F. HOFFMANN-LA ROCHE AG
(B) RUE: Grenzacherstrasse 124
(C) VILLE: Bâle
(D) ETAT: BS
(E) PAYS: Suisse
(F) CODE POSTAL: CH-4010
10 (G) TELEPHONE: 061-6884256
(H) TELECOPIEUR: 061-6881395
(I) TELEX: 962292/965542 hlr ch

- (II) TITRE DE L'INVENTION: Séquence d'ADN codant pour
une protéine de canal sodique, sa production et
15 son utilisation

(III) NOMBRE DES SEQUENCES: 5

(IV) FORME LISIBLE PAR ORDINATEUR:

- (A) TYPE DE SUPPORT: disque souple
(B) ORDINATEUR: IBM PC compatible
20 (C) SYSTEME D'EXPLOITATION: PC-DOS/MS-DOS
(D) LOGICIEL: PatentIn Release n° 1.0, version
1.30

(2) INFORMATIONS POUR SEQ ID n° 1:

(I) CARACTERISTIQUES DE LA SEQUENCE:

- 25 (A) LONGUEUR: 5 908 paires de bases
(B) TYPE: acide nucléique
(C) TYPE DE BRIN: simple
(D) TOPOLOGIE: linéaire

(II) TYPE DE MOLECULE: ADNc

30 (III) HYPOTHETIQUE: non

(IV) ANTISENS: non

(VI) SOURCE D'ORIGINE:

- (A) ORGANISME: rat
(B) TYPE DE TISSU: ganglions de la racine dor-
35 sale

(C) TYPE DE CELLULE: nerf périphérique

(XI) DESCRIPTION DE LA SEQUENCE: SEQ ID n° 1:

1 GAAGTCACAG GAGTGTCTGT CAGCGAGAGG AAGAAGGGAG AGTTTACTGA
 5 51 GTGTCTTCTG CCCCTCCTCA GGGTGAAGAT GGAGGAGAGG TACTACCCGG
 101 TGATCTTCCC GGACGAGCGG AATTTCCGCC CCTTCACTTC CGACTCTCTG
 151 GCTGCCATAG AGAAGCGGAT TGCTATCCAA AAGGAGAGGA AGAAGTCCAA
 201 AGACAAGGCG GCAGCTGAGC CCCAGCCTCG GCCTCAGCTT GACCTAAAGG
 10 251 CCTCCAGGAA GTTACCTAAG CTTTATGGTG ACATTCCCCC TGAGCTTGTA
 301 GCGAAGCCTC TGGAAGACCT GGACCCATTC TACAAAGACC ATAAGACATT
 351 CATGGTGTTG AACAAGAAGA GAACAATTTA TCGCTTCAGC GCCAAGCGGG
 401 CCTTGTTTAT TCTGGGGCCT TTTAATCCCC TCAGAAGCTT AATGATTTCG
 15 451 ATCTCTGTCC ATTCAGTCTT TAGCATGTTC ATCATCTGCA CGGTGATCAT
 501 CAACTGTATG TTCATGGCGA ATTCTATGGA GAGAAGTTTC GACAACGACA
 551 TTCCCGAATA CGTCTTCATT GGGATTTATA TTTTAGAAGC TGTGATTAAA
 20 601 ATATTGGCAA GAGGCTTCAT TGTGGATGAG TTTTCCTTCC TCCGAGATCC
 651 GTGGAACTGG CTGGACTTCA TTGTCATTGG AACAGCGATC GCAACTTGTT
 701 TTCCGGGCAG CCAAGTCAAT CTTTCAGCTC TTCGTACCTT CCGAGTGTTT
 751 AGAGCTCTGA AGGCGATTTT AGTTATCTCA GGTCTGAAGG TCATCGTAGG
 25 801 TGCCCTGCTG CGCTCGGTGA AGAAGCTGGT AGACGTGATG GTCCTCACTC
 851 TCTTCTGCCT CAGCATCTTT GCCCTGGTCG GTCAGCAGCT GTTCATGGGA
 901 ATTCTGAACC AGAAGTGTAT TAAGCACAAC TGTGGCCCCA ACCCTGCATC
 951 CAACAAGGAT TGCTTTGAAA AGGAAAAAGA TAGCGAAGAC TTCATAATGT
 30 1001 GTGGTACCTG GCTCGGCAGC AGACCCTGTC CCAATGGTTC TACGTGCGAT
 1051 AAAACCACAT TGAACCCAGA CAATAATTAT ACAAAGTTTG ACAACTTTGG
 1101 CTGGTCCTTT CTCGCCATGT TCCGGGTTAT GACTCAAGAC TCCTGGGAGA
 1151 GGCTTTACCG ACAGATCCTG CGGACCTCTG GGATCTACTT TGTCTTCTTC
 35

1201 TTCGTGGTGG TCATCTTCCT GGGCTCCTTC TACCTGCTTA ACCTAACCCCT
 1251 GGCTGTTGTC ACCATGGCTT ATGAAGAACA GAACAGAAAT GTAGCTGCTG
 1301 AGACAGAGGC CAAGGAGAAA ATGTTTCAGG AAGCCCAGCA GCTGTTAAGG
 5 1351 GAGGAGAAGG AGGCTCTGGT TGCCATGGGA ATTGACAGAA GTTCCCTTAA
 1401 TTCCCTTCAA GCTTCATCCT TTTCCCCGAA GAAGAGGAAG TTTTTCGGTA
 1451 GTAAGACAAG AAAGTCCTTC TTTATGAGAG GGTCCAAGAC GGCCCAAGCC
 1501 TCAGCGTCTG ATTCAGAGGA CGATGCCTCT AAAAATCCAC AGCTCCTTGA
 10 1551 GCAGACCAAA CGACTGTCCC AGAACTTGCC AGTGGATCTC TTTGATGAGC
 1601 ACGTGGACCC CCTCCACAGG CAGAGAGCGC TGAGCGCTGT CAGTATCTTA
 1651 ACCATCACCA TGCAGGAACA AGAAAAATTC CAGGAGCCTT GTTTCCCATG
 1701 TGGGAAAAAT TTGGCCTCTA AGTACCTGGT GTGGGACTGT AGCCCTCAGT
 15 1751 GGCTGTGCAT AAAGAAGGTC CTGCGGACCA TCATGACGGA TCCCTTTACT
 1801 GAGCTGGCCA TCACCATCTG CATCATCATC AATACCGTTT TCTTAGCCGT
 1851 GGAGCACCAC AACATGGATG ACAACTTAAA GACCATACTG AAAATAGGAA
 20 1901 ACTGGGTTTT CACGGGAATT TTCATAGCGG AAATGTGTCT CAAGATCATC
 1951 GCGCTCGACC CTTACCACTA CTTCCGGCAC GGCTGGAATG TTTTGTACAG
 2001 CATCGTGGCC CTCCTGAGTC TCGCTGATGT GCTCTACAAC AACTGTCTG
 2051 ATAACAATAG GTCTTTCTTG GCTTCCCTCA GAGTGCTGAG GGTCTTCAAG
 25 2101 TTAGCCAAAT CCTGGCCAC GTTAAACACT CTCATTAAGA TCATCGGCCA
 2151 CTCCGTGGGC GCGCTTGGA ACCTGACTGT GGTCCTGACT ATCGTGGTCT
 2201 TCATCTTTTC TGTGGTGGGC ATGCGGCTCT TCGGCACCAA GTTTAACAAG
 2251 ACCGCCTACG CCACCCAGGA GCGGCCAGG CGGCGCTGGC ACATGGATAA
 30 2301 TTTCTACCAC TCCTTCCTGG TGGTGTTCG CATCCTCTGT GGGGAATGGA
 2351 TCGAGAACAT GTGGGGCTGC ATGCAGGATA TGGACGGCTC CCCGTTGTGC
 2401 ATCATTGTCT TTGTCCTGAT AATGGTGATC GGGAAGCTTG TGGTGCTTAA
 35 2451 CCTCTTCATT GCCTTGCTGC TCAATTCCTT CAGCAATGAG GAGAAGGATG

2501 GGAGCCTGGA AGGAGAGACC AGGAAAACCA AAGTGCAGCT AGCCCTGGAT
2551 CGGTTCCGCC GGGCCTTCTC CTTCATGCTG CACGCTCTTC AGAGTTTTTG
2601 TTGCAAGAAA TGCAGGAGGA AAAACTCGCC AAAGCCAAAA GAGACAACAG
5 2651 AAAGCTTTGC TGGTGAGAAT AAAGACTCAA TCCTCCCGGA TGCGAGGCCC
2701 TGGAAGGAGT ATGATACAGA CATGGCTTTG TACTCTGGAC AGGCCGGGGC
2751 TCCGCTGGCC CCACTCGCAG AGGTAGAGGA CGATGTGGAA TATTGTGGTG
2801 AAGGCGGTGC CCTACCCACC TCACAACATA GTGCTGGAGT TCAGGCCGGT
10 2851 GACCTCCCTC CAGAGACCAA GCAGCTCACT AGCCCGGATG ACCAAGGGGT
2901 TGAAATGGAA GTATTTTCTG AAGAAGATCT GCATTTAAGC ATACAGAGTC
2951 CTCGAAAGAA GTCTGACGCA GTGAGCATGC TCTCGGAATG CAGCACAATT
3001 GACCTGAATG ATATCTTTAG AAATTTACAG AAAACAGTTT CCCCCAAAAA
15 3051 GCAGCCAGAT AGATGCTTTC CCAAGGGCCT TAGTTGTCAC TTTCTATGCC
3101 ACAAAACAGA CAAGAGAAAG TCCCCCTGGG TCCTGTGGTG GAACATTCGG
3151 AAAACCTGCT ACCAAATCGT GAAGCACAGC TGGTTTGAGA GTTTCATAAT
20 3201 CTTTGTATT CTGCTGAGCA GTGGAGCGCT GATATTTGAA GATGTCAATC
3251 TCCCCAGCCG GCCCCAAGTT GAGAAATTAC TAAGGTGTAC CGATAATATT
3301 TTCACATTTA TTTTCCTCCT GGAAATGATC CTGAAGTGGG TGGCCTTTGG
3351 ATTCCGGAGG TATTTACCA GTGCCTGGTG CTGGCTTGAT TTCCTCATTG
25 3401 TGGTGGTGTC TGTGCTCAGT CTCATGAATC TACCAAGCTT GAAGTCCTTC
3451 CGGACTCTGC GGGCCCTGAG ACCTCTGCGG GCGCTGTCCC AGTTTGAAGG
3501 AATGAAGGTT GTCGTCTACG CCCTGATCAG CGCCATACCT GCCATTCTCA
3551 ATGTCTTGCT GGTCTGCCTC ATTTTCTGGC TCGTATTTTG TATCTTGGGA
30 3601 GTAAATTTAT TTTCTGGGAA GTTTGAAGG TGCATTAACG GGACAGACAT
3651 AAATATGTAT TTGGATTTTA CCGAAGTTCC GAACCGAAGC CAATGTAACA
3701 TTAGTAATTA CTCGTGGAAG GTCCCGCAGG TCAACTTTGA CAACGTGGGG
35 3751 AATGCCTATC TCGCCCTGCT GCAAGTGGCA ACCTATAAGG GCTGGCTGGA

3801 AATCATGAAT GCTGCTGTCT ATTCCAGAGA GAAAGACGAG CAGCCGGACT
 3851 TTGAGGCGAA CCTCTACGCG TATCTCTACT TTGTGGTTTT TATCATCTTC
 3901 GGCTCCTTCT TTACCCTGAA CCTCTTTATC GGTGTTATTA TTGACAACTT
 5 3951 CAATCAGCAG CAGAAAAAGT TAGGTGGCCA AGACATTTTT ATGACAGAAG
 4001 AACAGAAGAA ATATTACAAT GCAATGAAAA AGTTAGGAAC CAAGAAACCT
 4051 CAAAAGCCCA TCCCAAGGCC CCTGAACAAA TGTCAAGCCT TTGTGTTCTGA
 4101 CCTGGTCACA AGCCAGGTCT TTGACGTCAT CATTCTGGGT CTTATTGTCT
 10 4151 TAAATATGAT TATCATGATG GCTGAATCTG CCGACCAGCC CAAAGATGTG
 4201 AAGAAAACCT TTGATATCCT CAACATAGCC TTCGTGGTCA TCTTTACCAT
 4251 AGAGTGTCTC ATCAAAGTCT TTGCTTTGAG GCAACACTAC TTCACCAATG
 15 4301 GCTGGAACCT ATTTGATTGT GTGGTCGTGG TTCTTTCTAT CATTAGTACC
 4351 CTGGTTTCCC GCTTGGAGGA CAGTGACATT TCTTTCCCGC CCACGCTCTT
 4401 CAGAGTCGTC CGCTTGGCTC GGATTGGTCG AATCCTCAGG CTGGTCCGGG
 4451 CTGCCCCGGG AATCAGGACC CTCCTCTTTG CTTTGATGAT GTCTCTCCCC
 20 4501 TCTCTCTTCA ACATCGGTCT GCTGCTCTTC CTGGTGATGT TCATTTACGC
 4551 CATCTTTGGG ATGAGCTGGT TTTCCAAAGT GAAGAAGGGC TCCGGGATCG
 4601 ACGACATCTT CAACTTCGAG ACCTTTACGG GCAGCATGCT GTGCCTCTTC
 4651 CAGATAACCA CTTCGGCTGG CTGGGATACC CTCCTCAACC CCATGCTGGA
 25 4701 GGCAAAAGAA CACTGCAACT CCTCCTCCCA AGACAGCTGT CAGCAGCCGC
 4751 AGATAGCCGT CGTCTACTTC GTCAGTTACA TCATCATCTC CTTCTCATC
 4801 GTGGTCAACA TGTACATCGC TGTGATCCTC GAGAACTTCA ACACAGCCAC
 4851 GGAGGAGAGC GAGGACCCTC TGGGAGAGGA CGACTTTGAA ATCTTCTATG
 30 4901 AGGTCTGGGA GAAGTTTGAC CCCGAGGCGT CGCAGTTCAT CCAGTATTCG
 4951 GCCCTCTCTG ACTTTGCGGA CGCCCTGCCG GAGCCGTTGC GTGTGGCCAA
 5001 GCCGAATAAG TTTCAGTTTC TAGTGATGGA CTTGCCCATG GTGATGGGCG
 35 5051 ACCGCCTCCA TTGCATGGAT GTTCTCTTTG CTTTCACTAC CAGGGTCCTC

- 46 -

5101 GGGGACTCCA GCGGCTTGGA TACCATGAAA ACCATGATGG AGGAGAAGTT
 5151 TATGGAGGCC AACCCTTTTA AGAAGCTCTA CGAGCCCATATA GTCACCACCA
 5201 CCAAGAGGAA GGAGGAGGAG CAAGGCGCCG CCGTCATCCA GAGGGCCTAC
 5 5251 CGGAAACACA TGGAGAAGAT GGTCAAACCTG AGGCTGAAGG ACAGGTCAAG
 5301 TTCATCGCAC CAGGTGTTTT GCAATGGAGA CTTGTCCAGC TTGGATGTGG
 5351 CCAAGGTCAA GGTTCAACAAT GACTGAACCC TCATCTCCAC CCCTACCTCA
 5401 CTGCCTCACA GCTTAGCCTC CAGCCTCTGG CGAGCAGGCG GCAGACTCAC
 10 5451 TGAACACAGG CCGTTCGATC TGTGTTTTTG GCTGAACGAG GTGACAGGTT
 5501 GCGCTCCATT TTTAAATGAC TCTTGAAAG ATTTTCATGTA GAGAGATGTT
 5551 AGAAGGGACT GCAAAGGACA CCGACCATAA CGGAAGGCCT GGAGGACAGT
 5601 CCAACTTACA TAAAGATGAG AAACAAGAAG GAAAGATCCC AGGAAAACCTT
 15 5651 CAGATTGTGT TCTCAGTACA TCCCCCAATG TGTCTGTTCG GTGTTTTGAG
 ----- 5701 TATGTGACCT GCCACATGTA GCTCTTTTTT GCATGTACGT CAAAACCCTG
 5751 CAGTAAGTTG ATAGCTTGCT ACGGGTGTTT CTACCAGCAT CACAGAATTG
 5801 GGTGTATGAC TCAAACCTAA AAGCATGACT CTGACTTGTC AGTCAGCACC
 20 5851 CCGACTTTCA GACGCTCCAA TCTCTGTCCC AGGTGTCTAA CGAATAAATA
 5901 GGTAAG

(3) INFORMATIONS POUR SEQ ID n° 2:

- 25 (I) CARACTERISTIQUES DE LA SEQUENCE:
 (A) LONGUEUR: 1 765 aminoacides
 (B) TYPE: aminoacide
 (C) TYPE DE BRIN: -
 (D) TOPOLOGIE: sans rapport
 (II) TYPE DE MOLECULE: protéine
 30 (III) HYPOTHETIQUE: oui
 (VI) SOURCE D'ORIGINE:
 (A) ORGANISME: rat
 (B) TYPE DE TISSU: ganglions de la racine dor-
 sale
 35 (C) TYPE DE CELLULE: nerf périphérique

- 47 -

(XI) DESCRIPTION DE LA SEQUENCE: SEQ ID n° 2:

Met Glu Glu Arg Tyr Tyr Pro Val Ile Phe Pro Asp Glu Arg Asn Phe
 1 5 10 15
 Arg Pro Phe Thr Ser Asp Ser Leu Ala Ala Ile Glu Lys Arg Ile Ala
 5 20 25 30
 Ile Gln Lys Glu Arg Lys Lys Ser Lys Asp Lys Ala Ala Ala Glu Pro
 35 40 45
 Gln Pro Arg Pro Gln Leu Asp Leu Lys Ala Ser Arg Lys Leu Pro Lys
 50 55 60
 10 Leu Tyr Gly Asp Ile Pro Pro Glu Leu Val Ala Lys Pro Leu Glu Asp
 65 70 75 80
 Leu Asp Pro Phe Tyr Lys Asp His Lys Thr Phe Met Val Leu Asn Lys
 85 90 95
 15 Lys Arg Thr Ile Tyr Arg Phe Ser Ala Lys Arg Ala Leu Phe Ile Leu
 100 105 110
 Gly Pro Phe Asn Pro Leu Arg Ser Leu Met Ile Arg Ile Ser Val His
 115 120 125
 Ser Val Phe Ser Met Phe Ile Ile Cys Thr Val Ile Ile Asn Cys Met
 130 135 140
 20 Phe Met Ala Asn Ser Met Glu Arg Ser Phe Asp Asn Asp Ile Pro Glu
 145 150 155 160
 Tyr Val Phe Ile Gly Ile Tyr Ile Leu Glu Ala Val Ile Lys Ile Leu
 165 170 175
 25 Ala Arg Gly Phe Ile Val Asp Glu Phe Ser Phe Leu Arg Asp Pro Trp
 180 185 190
 Asn Trp Leu Asp Phe Ile Val Ile Gly Thr Ala Ile Ala Thr Cys Phe
 195 200 205
 Pro Gly Ser Gln Val Asn Leu Ser Ala Leu Arg Thr Phe Arg Val Phe
 210 215 220
 30 Arg Ala Leu Lys Ala Ile Ser Val Ile Ser Gly Leu Lys Val Ile Val
 225 230 235 240
 Gly Ala Leu Leu Arg Ser Val Lys Lys Leu Val Asp Val Met Val Leu
 245 250 255
 35 Thr Leu Phe Cys Leu Ser Ile Phe Ala Leu Val Gly Gln Gln Leu Phe
 260 265 270

- 48 -

Met Gly Ile Leu Asn Gln Lys Cys Ile Lys His Asn Cys Gly Pro Asn
 275 280 285
 Pro Ala Ser Asn Lys Asp Cys Phe Glu Lys Glu Lys Asp Ser Glu Asp
 290 295 300
 5 Phe Ile Met Cys Gly Thr Trp Leu Gly Ser Arg Pro Cys Pro Asn Gly
 305 310 315 320
 Ser Thr Cys Asp Lys Thr Thr Leu Asn Pro Asp Asn Asn Tyr Thr Lys
 325 330 335
 Phe Asp Asn Phe Gly Trp Ser Phe Leu Ala Met Phe Arg Val Met Thr
 10 340 345 350
 Gln Asp Ser Trp Glu Arg Leu Tyr Arg Gln Ile Leu Arg Thr Ser Gly
 355 360 365
 Ile Tyr Phe Val Phe Phe Phe Val Val Val Ile Phe Leu Gly Ser Phe
 370 375 380
 15 Tyr Leu Leu Asn Leu Thr Leu Ala Val Val Thr Met Ala Tyr Glu Glu
 385 390 395 400
 Gln Asn Arg Asn Val Ala Ala Glu Thr Glu Ala Lys Glu Lys Met Phe
 405 410 415
 Gln Glu Ala Gln Gln Leu Leu Arg Glu Glu Lys Glu Ala Leu Val Ala
 20 420 425 430
 Met Gly Ile Asp Arg Ser Ser Leu Asn Ser Leu Gln Ala Ser Ser Phe
 435 440 445
 Ser Pro Lys Lys Arg Lys Phe Phe Gly Ser Lys Thr Arg Lys Ser Phe
 450 455 460
 25 Phe Met Arg Gly Ser Lys Thr Ala Gln Ala Ser Ala Ser Asp Ser Glu
 465 470 475 480
 Asp Asp Ala Ser Lys Asn Pro Gln Leu Leu Glu Gln Thr Lys Arg Leu
 485 490 495
 Ser Gln Asn Leu Pro Val Asp Leu Phe Asp Glu His Val Asp Pro Leu
 30 500 505 510
 His Arg Gln Arg Ala Leu Ser Ala Val Ser Ile Leu Thr Ile Thr Met
 515 520 525
 Gln Glu Gln Glu Lys Phe Gln Glu Pro Cys Phe Pro Cys Gly Lys Asn
 530 535 540

- 49 -

	Leu	Ala	Ser	Lys	Tyr	Leu	Val	Trp	Asp	Cys	Ser	Pro	Gln	Trp	Leu	Cys
	545					550					555					560
	Ile	Lys	Lys	Val	Leu	Arg	Thr	Ile	Met	Thr	Asp	Pro	Phe	Thr	Glu	Leu
						565					570					575
5	Ala	Ile	Thr	Ile	Cys	Ile	Ile	Ile	Asn	Thr	Val	Phe	Leu	Ala	Val	Glu
						580					585					590
	His	His	Asn	Met	Asp	Asp	Asn	Leu	Lys	Thr	Ile	Leu	Lys	Ile	Gly	Asn
						595					600					605
	Trp	Val	Phe	Thr	Gly	Ile	Phe	Ile	Ala	Glu	Met	Cys	Leu	Lys	Ile	Ile
10		610					615					620				
	Ala	Leu	Asp	Pro	Tyr	His	Tyr	Phe	Arg	His	Gly	Trp	Asn	Val	Phe	Asp
	625					630					635					640
	Ser	Ile	Val	Ala	Leu	Leu	Ser	Leu	Ala	Asp	Val	Leu	Tyr	Asn	Thr	Leu
						645					650					655
15	Ser	Asp	Asn	Asn	Arg	Ser	Phe	Leu	Ala	Ser	Leu	Arg	Val	Leu	Arg	Val
						660					665					670
	Phe	Lys	Leu	Ala	Lys	Ser	Trp	Pro	Thr	Leu	Asn	Thr	Leu	Ile	Lys	Ile
						675					680					685
	Ile	Gly	His	Ser	Val	Gly	Ala	Leu	Gly	Asn	Leu	Thr	Val	Val	Leu	Thr
20		690					695					700				
	Ile	Val	Val	Phe	Ile	Phe	Ser	Val	Val	Gly	Met	Arg	Leu	Phe	Gly	Thr
	705					710					715					720
	Lys	Phe	Asn	Lys	Thr	Ala	Tyr	Ala	Thr	Gln	Glu	Arg	Pro	Arg	Arg	Arg
						725					730					735
25	Trp	His	Met	Asp	Asn	Phe	Tyr	His	Ser	Phe	Leu	Val	Val	Phe	Arg	Ile
						740					745					750
	Leu	Cys	Gly	Glu	Trp	Ile	Glu	Asn	Met	Trp	Gly	Cys	Met	Gln	Asp	Met
						755					760					765
30	Asp	Gly	Ser	Pro	Leu	Cys	Ile	Ile	Val	Phe	Val	Leu	Ile	Met	Val	Ile
		770					775					780				
	Gly	Lys	Leu	Val	Val	Leu	Asn	Leu	Phe	Ile	Ala	Leu	Leu	Leu	Asn	Ser
	785					790					795					800
	Phe	Ser	Asn	Glu	Glu	Lys	Asp	Gly	Ser	Leu	Glu	Gly	Glu	Thr	Arg	Lys
35						805					810					815

- 50 -

Thr Lys Val Gln Leu Ala Leu Asp Arg Phe Arg Arg Ala Phe Ser Phe
 820 825 830
 Met Leu His Ala Leu Gln Ser Phe Cys Cys Lys Lys Cys Arg Arg Lys
 835 840 845
 5 Asn Ser Pro Lys Pro Lys Glu Thr Thr Glu Ser Phe Ala Gly Glu Asn
 850 855 860
 Lys Asp Ser Ile Leu Pro Asp Ala Arg Pro Trp Lys Glu Tyr Asp Thr
 865 870 875 880
 Asp Met Ala Leu Tyr Thr Gly Gln Ala Gly Ala Pro Leu Ala Pro Leu
 885 890 895
 10 Ala Glu Val Glu Asp Asp Val Glu Tyr Cys Gly Glu Gly Gly Ala Leu
 900 905 910
 Pro Thr Ser Gln His Ser Ala Gly Val Gln Ala Gly Asp Leu Pro Pro
 915 920 925
 15 Glu Thr Lys Gln Leu Thr Ser Pro Asp Asp Gln Gly Val Glu Met Glu
 930 935 940
 Val Phe Ser Glu Glu Asp Leu His Leu Ser Ile Gln Ser Pro Arg Lys
 945 950 955 960
 Lys Ser Asp Ala Val Ser Met Leu Ser Glu Cys Ser Thr Ile Asp Leu
 965 970 975
 20 Asn Asp Ile Phe Arg Asn Leu Gln Lys Thr Val Ser Pro Lys Lys Gln
 980 985 990
 Pro Asp Arg Cys Phe Pro Lys Gly Leu Ser Cys His Phe Leu Cys His
 995 1000 1005
 25 Lys Thr Asp Lys Arg Lys Ser Pro Trp Val Leu Trp Trp Asn Ile Arg
 1010 1015 1020
 Lys Thr Cys Tyr Gln Ile Val Lys His Ser Trp Phe Glu Ser Phe Ile
 1025 1030 1035 1040
 Ile Phe Val Ile Leu Leu Ser Ser Gly Ala Leu Ile Phe Glu Asp Val
 1045 1050 1055
 30 Asn Leu Pro Ser Arg Pro Gln Val Glu Lys Leu Leu Arg Cys Thr Asp
 1060 1065 1070
 Asn Ile Phe Thr Phe Ile Phe Leu Leu Glu Met Ile Leu Lys Trp Val
 1075 1080 1085

- 51 -

	Ala Phe Gly Phe Arg Arg Tyr Phe Thr Ser Ala Trp Cys Trp Leu Asp
	1090 1095 1100
	Phe Leu Ile Val Val Val Ser Val Leu Ser Leu Met Asn Leu Pro Ser
	1105 1110 1115 1120
5	Leu Lys Ser Phe Arg Thr Leu Arg Ala Leu Arg Pro Leu Arg Ala Leu
	1125 1130 1135
	Ser Gln Phe Glu Gly Met Lys Val Val Val Tyr Ala Leu Ile Ser Ala
	1140 1145 1150
	Ile Pro Ala Ile Leu Asn Val Leu Leu Val Cys Leu Ile Phe Trp Leu
	1155 1160 1165
10	Val Phe Cys Ile Leu Gly Val Asn Leu Phe Ser Gly Lys Phe Gly Arg
	1170 1175 1180
	Cys Ile Asn Gly Thr Asp Ile Asn Met Tyr Leu Asp Phe Thr Glu Val
	1185 1190 1195 1200
15	Pro Asn Arg Ser Gln Cys Asn Ile Ser Asn Tyr Ser Trp Lys Val Pro
	1205 1210 1215
	Gln Val Asn Phe Asp Asn Val Gly Asn Ala Tyr Leu Ala Leu Leu Gln
	1220 1225 1230
	Val Ala Thr Tyr Lys Gly Trp Leu Glu Ile Met Asn Ala Ala Val Asp
	1235 1240 1245
20	Ser Arg Glu Lys Asp Glu Gln Pro Asp Phe Glu Ala Asn Leu Tyr Ala
	1250 1255 1260
	Tyr Leu Tyr Phe Val Val Phe Ile Ile Phe Gly Ser Phe Phe Thr Leu
	1265 1270 1275 1280
25	Asn Leu Phe Ile Gly Val Ile Ile Asp Asn Phe Asn Gln Gln Gln Lys
	1285 1290 1295
	Lys Leu Gly Gly Gln Asp Ile Phe Met Thr Glu Glu Gln Lys Lys Tyr
	1300 1305 1310
	Tyr Asn Ala Met Lys Lys Leu Gly Thr Lys Lys Pro Gln Lys Pro Ile
30	1315 1320 1325
	Pro Arg Pro Leu Asn Lys Cys Gln Ala Phe Val Phe Asp Leu Val Thr
	1330 1335 1340
	Ser Gln Val Phe Asp Val Ile Ile Leu Gly Leu Ile Val Leu Asn Met
	1345 1350 1355 1360

- 52 -

Ile Ile Met Met Ala Glu Ser Ala Asp Gln Pro Lys Asp Val Lys Lys
 1365 1370 1375
 Thr Phe Asp Ile Leu Asn Ile Ala Phe Val Val Ile Phe Thr Ile Glu
 1380 1385 1390
 5 Cys Leu Ile Lys Val Phe Ala Leu Arg Gln His Tyr Phe Thr Asn Gly
 1395 1400 1405
 Trp Asn Leu Phe Asp Cys Val Val Val Val Leu Ser Ile Ile Ser Thr
 1410 1415 1420
 Leu Val Ser Arg Leu Glu Asp Ser Asp Ile Ser Phe Pro Pro Thr Leu
 10 1425 1430 1435 1440
 Phe Arg Val Val Arg Leu Ala Arg Ile Gly Arg Ile Leu Arg Leu Val
 1445 1450 1455
 Arg Ala Ala Arg Gly Ile Arg Thr Leu Leu Phe Ala Leu Met Met Ser
 1460 1465 1470
 15 Leu Pro Ser Leu Phe Asn Ile Gly Leu Leu Leu Phe Leu Val Met Phe
 1475 1480 1485
 Ile Tyr Ala Ile Phe Gly Met Ser Trp Phe Ser Lys Val Lys Lys Gly
 1490 1495 1500
 Ser Gly Ile Asp Asp Ile Phe Asn Phe Glu Thr Phe Thr Gly Ser Met
 20 1505 1510 1515 1520
 Leu Cys Leu Phe Gln Ile Thr Thr Ser Ala Gly Trp Asp Thr Leu Leu
 1525 1530 1535
 Asn Pro Met Leu Glu Ala Lys Glu His Cys Asn Ser Ser Ser Gln Asp
 1540 1545 1550
 25 Ser Cys Gln Gln Pro Gln Ile Ala Val Val Tyr Phe Val Ser Tyr Ile
 1555 1560 1565
 Ile Ile Ser Phe Leu Ile Val Val Asn Met Tyr Ile Ala Val Ile Leu
 1570 1575 1580
 Glu Asn Phe Asn Thr Ala Thr Glu Glu Ser Glu Asp Pro Leu Gly Glu
 30 1585 1590 1595 1600
 Asp Asp Phe Glu Ile Phe Tyr Glu Val Trp Glu Lys Phe Asp Pro Glu
 1605 1610 1615
 Ala Ser Gln Phe Ile Gln Tyr Ser Ala Leu Ser Asp Phe Ala Asp Ala
 1620 1625 1630
 35 Leu Pro Glu Pro Leu Arg Val Ala Lys Pro Asn Lys Phe Gln Phe Leu

- 53 -

	1635	1640	1645
	Val Met Asp Leu Pro Met Val Met Gly Asp Arg Leu His Cys Met Asp		
	1650	1655	1660
5	Val Leu Phe Ala Phe Thr Thr Arg Val Leu Gly Asp Ser Ser Gly Leu		
	1665	1670	1675 1680
	Asp Thr Met Lys Thr Met Met Glu Glu Lys Phe Met Glu Ala Asn Pro		
	1685	1690	1695
	Phe Lys Lys Leu Tyr Glu Pro Ile Val Thr Thr Thr Lys Arg Lys Glu		
10	1700	1705	1710
	Glu Glu Gln Gly Ala Ala Val Ile Gln Arg Ala Tyr Arg Lys His Met		
	1715	1720	1725
	Glu Lys Met Val Lys Leu Arg Leu Lys Asp Arg Ser Ser Ser Ser His		
	1730	1735	1740
15	Gln Val Phe Cys Asn Gly Asp Leu Ser Ser Leu Asp Val Ala Lys Val		
	1745	1750	1755 1760
	Lys Val His Asn Asp		
	1765		

(4) INFORMATIONS POUR SEQ ID n° 3:

- 20 (I) CARACTERISTIQUES DE LA SEQUENCE:
- (A) LONGUEUR: 856 paires de bases
 - (B) TYPE: acide nucléique
 - (C) TYPE DE BRIN: simple
 - (D) TOPOLOGIE: linéaire
- 25 (II) TYPE DE MOLECULE: ADNc
- (III) HYPOTHETIQUE: non
- (IV) ANTISENS: non
- (VI) SOURCE D'ORIGINE:
- (A) ORGANISME: humain
- 30 (B) TYPE DE TISSU: ganglions de la racine dorsale
- (C) TYPE DE CELLULE: nerf périphérique
- (XI) DESCRIPTION DE LA SEQUENCE: SEQ ID n° 3:

35 ¹ GCTGAGCAGT GGGGCACTGA TATTTGAAGA TGTTACCTT GAGAACCAAC

- 54 -

51 CCAAATCCA AGAATTACTA AATTGTACTG ACATTATTTT TACACATATT
 101 TTTATCCTGG AGATGGTACT AAAATGGGTA GCCTTCGGAT TTGGAAAGTA
 151 TTTCACCAGT GCCTGGTGCT GCCTTGATTT CATCATTGTG ATTGTCTCTG
 5 201 TGACCACCCT CATTAACCTA ATGGAATTGA AGTCCTCCG GACTCTACGA
 251 GCACTGAGGC CTCTTCGTGC GCTGTCCCAG TTTGAAGGAA TGAAGGTGGT
 301 GGTCAATGCT CTCATAGGTG CCATACCTGC CATTCTGAAT GTTTTGCTTG
 351 TCTGCCTCAT TTTCTGGCTC GTATTTTGTA TTCTGGGAGT ATACTTCTTT
 10 401 TCTGGAAAAT TTGGGAAATG CATTAATGGA ACAGACTCAG TTATAAATTA
 451 TACCATCATT ACAAATAAAA GTCAATGTGA AAGTGGCAAT TTCTCTTGGA
 501 TCAACCAGAA AGTCAACTTT GACAATGTGG GAAATGCTTA CCTCGCTCTG
 15 551 CTGCAAGTGG CAACATTTAA GGGCTGGATG GATATTATAT ATGCAGCTGT
 601 TGATTCCACA GAGAAAGAAC AACAGCCAGA GTTTGAGAGC AATTCACCTCG
 651 GTTACATTTA CTTCGTAGTC TTTATCATCT TTGGCTCATT CTTCACTCTG
 701 AATCTCTTCA TTGGCGTTAT CATTGACAAC TTCAACCAAC AGCAGAAAAA
 20 751 GTTAGGTGGC CAAGACATTT TTATGACAGA AGAACAGAAG AAATACTATA
 801 ATGCAATGAA AAAATTAGGA TCCAAAAAAC CTCAAAAACC CATTCCACGG
 851 CCCGTT

25

(5) INFORMATIONS POUR SEQ ID n° 4:

(I) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 702 paires de bases

(B) TYPE: acide nucléique

30 (C) TYPE DE BRIN: simple

(D) TOPOLOGIE: linéaire

(II) TYPE DE MOLECULE: RT-PCR

 (A) DESCRIPTION: /desc = "sonde d'ADN/
 domaine IV"

35

(III) HYPOTHETIQUE: non

- 55 -

(IV) ANTISENS: non

(VI) SOURCE D'ORIGINE:

(A) ORGANISME: rat

(B) TYPE DE TISSU: ganglions de la racine dorsale

(C) TYPE DE CELLULE: nerf périphérique

(XI) DESCRIPTION DE LA SEQUENCE: SEQ ID n° 4:

1 CTCAACATGG TTACGATGAT GGTGGAGACC GACGAGCAGG GCGAGGAGAA
 51 GACGAAGGTT CTGGGCAGAA TCAACCAGTT CTTTGTGGCC GTCTTCACGG
 101 GCGAGTGTGT GATGAAGATG TTCGCCCTGC GACAGTACTA TTTCACCAAC
 151 GGCTGGAACG TGTTGCACTT CATAGTGGTG ATCCTGTCCA TTGGGAGTCT
 201 GCTGTTTCT GCAATCCTTA AGTCACTGGA AAATACTTC TCCCCGACGC
 251 TCTTCCGGGT CATCCGTCTG GCCAGGATCG GCCGCATCCT CAGGCTGATC
 301 CGAGCAGCCA AGGGGATTCG CACGCTGCTC TTCGCCCTCA TGATGTCCCT
 351 GCCCGCCCTC TTCAACATCG GCCTCCTCCT CTTCCTCGTC ATGTTTCATCT
 401 ACTCCATCTT CGGCATGGCC AGCTTCGCTA ACGTCGTGGA CGAGGCCGGC
 451 ATCGACGACA TGTTCAACTT CAAGACCTTT GGCAACAGCA TGCTGTGCCT
 501 GTTCCAGATC ACCACCTCGG CCGGCTGGGA CGGCCTCCTC AGCCCCATCC
 551 TCAACACGGG GCCTCCCTAC TGCGACCCCA ACCTGCCCAA CAGCAACGGC
 601 TCCCGGGGGA ACTGCGGGAG CCCGGCGGTG GGCATCATCT TCTTCACCAC
 651 CTACATCATC ATCTCCTTCC TCATCGTGGT CAACATGTAT ATCGCAGTCA
 701 TC

(6) INFORMATIONS POUR SEQ ID n° 5:

25 (I) CARACTERISTIQUES DE LA SEQUENCE:
 (A) LONGUEUR: 5 334 paires de bases
 (B) TYPE: acide nucléique
 (C) TYPE DE BRIN: simple
 (D) TOPOLOGIE: linéaire
 30 (II) TYPE DE MOLECULE: RT-PCR
 (A) DESCRIPTION: ADNc
 (III) HYPOTHETIQUE: non
 (IV) ANTISENS: non
 (VI) SOURCE D'ORIGINE:
 35 (A) ORGANISME:

(B) TYPE DE TISSU:

(C) TYPE DE CELLULE:

(XI) DESCRIPTION DE LA SEQUENCE: SEQ ID n° 5:

5 1 GTCGACTCTA GATCAGGGTG AAGATGGAGG AGAGGTACTA CCCGGTGATC
 51 TTCCCGGACG AGCGGAATTT CCGCCCCCTTC ACTTCCGACT CTCTGGCTGC
 101 CATAGAGAAG CGGATTGCTA TCCAAAAGGA GAGGAAGAAG TCCAAAGACA
 151 AGGCGGCAGC TGAGCCCCAG CCTCGGCCTC AGCTTGACCT AAAGGCCTCC
10 201 AGGAAGTTAC CTAAGCTTTA TGGTGACATT CCCCCTGAGC TTGTAGCGAA
 251 GCCTCTGGAA GACCTGGACC CATTCTACAA AGACCATAAG ACATTCATGG
 301 TGTGTAACAA GAAGAGAACA ATTTATCGCT TCAGCGCCAA GCGGGCCTTG
 351 TTCATTCTGG GGCCTTTTAA TCCCCTCAGA AGCTTAATGA TTCGTATCTC
15 401 TGTCCATTCA GTCTTTAGCA TGTTTCATCAT CTGCACGGTG ATCATCAACT
 451 GTATGTTCAT GGCGAATTCT ATGGAGAGAA GTTTCGACAA CGACATTCCC
 501 GAATACGTCT TCATTGGGAT TTATATTTTA GAAGCTGTGA TAAAATATT
 551 GGCAAGAGGC TTCATTGTGG ATGAGTTTTC CTCCTCCGA GATCCGTGGA
20 601 ACTGGCTGGA CTTCAATTGTC ATTGGAACAG CGATCGCAAC TTGTTTTCCG
 651 GGCAGCCAAG TCAATCTTTC AGCTCTTCGT ACCTTCCGAG TGTTCAGAGC
 701 TCTGAAGGCG ATTTTCAGTTA TCTCAGGTCT GAAGGTCATC GTAGGTGCCC
 751 TGCTGCGCTC GGTGAAGAAG CTGGTAGACG TGATGGTCCT CACTCTCTTC
25 801 TGCCTCAGCA TCTTTGCCCT GGTGCGTCAG CAGCTGTTCA TGGGAATTCT
 851 GAACCAGAAG TGTATTAAGC ACAACTGTGG CCCCAACCCT GCATCCAACA
 901 AGGATTGCTT TGAAAAGGAA AAAGATAGCG AAGACTTCAT AATGTGTGGT
30 951 ACCTGGCTCG GCAGCAGACC CTGTCCCAAT GGTTCACGT GCGATAAAAC

1001 CACATTGAAC CCAGACAATA ATTATACAAA GTTTGACAAC TTTGGCTGGT
1051 CCTTTCTCGC CATGTTCCGG GTTATGACTC AAGACTCCTG GGAGAGGCTT
1101 TACCGACAGA TCCTGCGGAC CTCTGGGATC TACTTTGTCT TCTTCTTCGT
1151 GGTGGTCATC TTCCTGGGCT CCTTCTACCT GCTTAACCTA ACCCTGGCTG
1201 TTGTCACCAT GGCTTATGAA GAACAGAACA GAAATGTAGC TGCTGAGACA
1251 GAGGCCAAGG AGAAAATGTT TCAGGAAGCC CAGCAGCTGT TAAGGGAGGA
1301 GAAGGAGGCT CTGGTTGCCA TGGGAATTGA CAGAAGTTCC CTTAATTCCC
1351 TTCAAGCTTC ATCCTTTTCC CCGAAGAAGA GGAAGTTTTT CGGTAGTAAG
1401 ACAAGAAAGT CCTTCTTTAT GAGAGGGTCC AAGACGGCCC AAGCCTCAGC
1451 GTCTGATTCA GAGGACGATG CCTCTAAAAA TCCACAGCTC CTTGAGCAGA
1501 CCAAACGACT GTCCCAGAAC TTGCCAGTGG ATCTCTTTGA TGAGCACGTG
1551 GACCCCTCC ACAGGCAGAG AGCGCTGAGC GCTGTCAGTA TCTTAACCAT
1601 CACCATGCAG GAACAAGAAA AATTCCAGGA GCCTTGTTTC CCATGTGGGA
1651 AAAATTTGGC CTCTAAGTAC CTGGTGTGGG ACTGTAGCCC TCAGTGGCTG
1701 TGCATAAAGA AGGTCCTGCG GACCATCATG ACGGATCCCT TTACTGAGCT
1751 GGCCATCACC ATCTGCATCA TCATCAATAC CGTTTTCTTA GCCGTGGAGC
1801 ACCACAACAT GGATGACAAC TTAAAGACCA TACTGAAAAT AGGAAACTGG
1851 GTTTTCACGG GAATTTTCAT AGCGGAAATG TGTCTCAAGA TCATCGCGCT
1901 CGACCCTTAC CACTACTTCC GGCACGGCTG GAATGTTTTT GACAGCATCG
1951 TGGCCCTCCT GAGTCTCGCT GATGTGCTCT ACAACACACT GTCTGATAAC
2001 AATAGGTCTT TCTTGGCTTC CCTCAGAGTG CTGAGGGTCT TCAAGTTAGC
2051 CAAATCCTGG CCCACGTAA ACACTCTCAT TAAGATCATC GGCCACTCCG
2101 TGGGCGCGCT TGGAAACCTG ACTGTGGTCC TGACTATCGT GGTCTTCATC
2151 TTTTCTGTGG TGGGCATGCG GCTCTTCGGC ACCAAGTTTA ACAAGACCGC
2201 CTACGCCACC CAGGAGCGGC CCAGGCGGCG CTGGCACATG GATAATTTCT
2251 ACCACTCCTT CCTGGTGGTG TTCCGCATCC TCTGTGGGGA ATGGATCGAG
2301 AACATGTGGG GCTGCATGCA GGATATGGAC GGCTCCCCGT TGTGCATCAT
2351 TGTCTTTGTC CTGATAATGG TGATCGGGAA GCTTGTGGTG CTTAACCTCT

2401 TCATTGCCTT GCTGCTCAAT TCCTTCAGCA ATGAGGAGAA GGATGGGAGC
2451 CTGGAAGGAG AGACCAGGAA AACCAAAGTG CAGCTAGCCC TGGATCGGTT
2501 CCGCCGGGCC TTCTCCTTCA TGCTGCACGC TCTTCAGAGT TTTTGTTGCA
2551 AGAAATGCAG GAGGAAAAAC TCGCCAAAGC CAAAAGAGAC AACAGAAAGC
2601 TTTGCTGGTG AGAATAAAGA CTCAATCCTC CCGGATGCGA GGCCCTGGAA
2651 GGAGTATGAT ACAGACATGG CTTTGTACAC TGGACAGGCC GGGGCTCCGC
2701 TGGCCCCACT CGCAGAGGTA GAGGACGATG TGAATATTG TGGTGAAGGC
2751 GGTGCCCTAC CCACCTCACA ACATAGTGCT GGAGTTCAGG CCGGTGACCT
2801 CCCTCCAGAG ACCAAGCAGC TCACTAGCCC GGATGACCAA GGGGTTGAAA
2851 TGGAAGTATT TTCTGAAGAA GATCTGCATT TAAGCATACA GAGTCCTCGA
2901 AAGAAGTCTG ACGCAGTGAG CATGCTCTCG GAATGCAGCA CAATTGACCT
2951 GAATGATATC TTTAGAAATT TACAGAAAAC AGTTTCCCCC AAAAAGCAGC
3001 CAGATAGATG CTTTCCCAAG GGCCTTAGTT GTCACTTTCT ATGCCACAAA
3051 ACAGACAAGA GAAAGTCCCC CTGGGTCCCTG TGGTGGAACA TTCGGAAAAC
3101 CTGCTACCAA ATCGTGAAGC ACAGCTGGTT TGAGAGTTTC ATAATCTTTG
3151 TTATTCTGCT GAGCAGTGGA GCGCTGATAT TTGAAGATGT CAATCTCCCC
3201 AGCCGGCCCC AAGTTGAGAA ATTACTAAGG TGTACCGATA ATATTTTCAC
3251 ATTTATTTTC CTCCTGGAAG TGATCCTGAA GTGGGTGGCC TTTGGATTCC
3301 GGAGGTATTT CACCAGTGCC TGGTGCTGGC TTGATTTCTT CATTTGTGGT
2251 GTGTCTGTGC TCAGTCTCAT GAATCTACCA AGCTTGAAGT CCTTCCGGAC
3401 TCTGCGGGCC CTGAGACCTC TCGGGCGCT GTCCCAGTTT GAAGGAATGA
3451 AGGTTGTCTG CTACGCCCTG ATCAGCGCCA TACCTGCCAT TCTCAATGTC
3501 TTGCTGGTCT GCCTCATTTT CTGGCTCGTA TTTTGTATCT TGGGAGTAAA
3551 TTTATTTTCT GGAAGTTTG GAAGGTGCAT TAACGGGACA GACATAAATA
3601 TGTATTTGGA TTTTACCGAA GTTCCGAACC GAAGCCAATG TAACATTAGT
3651 AATTACTCGT GGAAGGTCCC GCAGGTCAAC TTTGACAACG TGGGGAATGC
3701 CTATCTCGCC CTGCTGCAAG TGGCAACCTA TAAGGGCTGG CTGGAAATCA
3751 TGAATGCTGC TGTCGATTCC AGAGAGAAAG ACGAGCAGCC GGACTTTGAG

3801 GCGAACCTCT ACGCGTATCT CTACTTTGTG GTTTTTATCA TCTTCGGCTC
 3851 CTTCTTTACC CTGAACCTCT TTATCGGTGT TATTATTGAC AACTTCAATC
 3901 AGCAGCAGAA AAAGTTAGGT GGCCAAGACA TCTTCATGAC TGAGGAGCAG
 3951 AAGAAATATT ACAATGCAAT GAAAAAGTTA GGAACCAAGA AACCTCAAAA
 4001 GCCCATCCCA AGGCCCCCTGA ACAAATGTCA AGCCTTTGTG TTCGACCTGG
 4051 TCACAAGCCA GGTCTTTGAC GTCATCATTG TGGGTCTTAT TGTCTTAAAT
 4101 ATGATTATCA TGATGGCTGA ATCTGCCGAC CAGCCCAAAG ATGTGAAGAA
 4151 AACCTTTGAT ATCCTCAACA TAGCCTTCGT GGTCACTTTT ACCATAGAGT
 4201 GTCTCATCAA AGTCTTTGCT TTGAGGCAAC ACTACTTCAC CAATGGCTGG
 4251 AACTTATTTG ATTGTGTGGT CGTGGTTCTT TCTATCATTG GTACCCTGGT
 4301 TTCCCGCTTG GAGGACAGTG ACATTTCTTT CCCGCCACG CTCTTCAGAG
 4351 TCGTCCGCTT GGCTCGGATT GGTGGAATCC TCAGGCTGGT CCGGGCTGCC
 4401 CGGGGAATCA GGACCCTCCT CTTTGCTTTG ATGATGTCTC TCCCCTCTCT
 4451 CTTCAACATC GGTCTGCTGC TCTTCCTGGT GATGTTCATT TACGCCATCT
 4501 TTGGGATGAG CTGGTTTTC AAAGTGAAGA AGGGCTCCGG GATCGACGAC
 4551 ATCTTCAACT TCGAGACCTT TACGGGCAGC ATGCTGTGCC TCTTCCAGAT
 4601 AACCCTTCG GCTGGCTGGG ATACCCTCCT CAACCCCATG CTGGAGGCAA
 4651 AAGAACACTG CAACTCCTCC TCCAAGACA GCTGTCAGCA GCCGCAGATA
 4701 GCCGTCGTCT ACTTCGTCAG TTACATCATC ATCTCCTTCC TCATCGTGGT
 4751 CAACATGTAC ATCGCTGTGA TCCTCGAGAA CTTCAACACA GCCACGGAGG
 4801 AGAGCGAGGA CCCTCTGGGA GAGGACGACT TTGAAATCTT CTATGAGGTC
 4851 TGGGAGAAGT TTGACCCCGA GGCGTCGCAG TTCATCCAGT ATTCGGCCCT
 4901 CTCTGACTTT GCGGACGCCC TGCCGGAGCC GTTGCCTGTG GCCAAGCCGA
 4951 ATAAGTTTCA GTTTCTAGTG ATGGACTTGC CCATGGTGAT GGGCGACCGC
 5001 CTCCATTGCA TGGATGTTCT CTTTGCTTTC ACTACCAGGG TCCTCGGGGA
 5051 CTCCAGCGGC TTGGATACCA TGAAAACCAT GATGGAGGAG AAGTTTATGG
 5101 AGGCCAACCC TTTTAAGAAG CTCTACGAGC CCATAGTCAC CACCACCAAG
 5151 AGGAAGGAGG AGGAGCAAGG CGCCGCCGTC ATCCAGAGGG CCTACCGGAA

5201 ACACATGGAG AAGATGGTCA AACTGAGGCT GAAGGACAGG TCAAGTTCAT
5251 CGCACCAGGT GTTTTGCAAT GGAGACTTGT CCAGCTTGGA TGTGGCCAAG
5301 GTCAAGGTTC ACAATGACTG AACCTCATC TAGA

REVENDICATIONS

1. Séquence d'ADN isolée, caractérisée en ce qu'elle comprend la séquence nucléotidique indiquée dans SEQ ID n° 1 et SEQ ID n° 3.
- 5 2. ADN selon la revendication 1, caractérisé en ce que ladite séquence d'ADN code pour une protéine de canal sodique ou un fragment de celle-ci.
3. ADN selon la revendication 2, caractérisé en ce que ladite protéine de canal sodique est la sous-unité α
10 ou un fragment de celle-ci.
4. ADN selon la revendication 3, caractérisé en ce que ladite protéine de canal sodique est résistante à la tétrodotoxine.
5. ADN selon la revendication 3 ou 4, caractérisé
15 en ce que ladite protéine de canal sodique se rencontre chez des mammifères.
6. ADN selon la revendication 3 ou 4, caractérisé en ce que ladite protéine de canal sodique se rencontre chez le rat.
- 20 7. ADN selon la revendication 3 ou 4, caractérisé en ce que ladite protéine de canal sodique se rencontre chez l'homme.
8. ADN selon la revendication 1, caractérisé en ce que ledit ADN est de l'ADNc.
- 25 9. ADN selon la revendication 1, caractérisé en ce que ledit ADN est de l'ADN synthétique.
10. Vecteurs d'expression, caractérisés en ce qu'ils comprennent l'ADN de la revendication 8.
11. Vecteurs d'expression, caractérisés en ce
30 qu'ils comprennent l'ADN synthétique de la revendication 9.
12. Cellules hôtes transformées par les vecteurs d'expression de la revendication 10.
13. Cellules hôtes transformées par les vecteurs
35 d'expression de la revendication 11.

14. Polynucléotide recombinant, caractérisé en ce qu'il comprend une séquence d'acide nucléique provenant de la séquence d'ADN de la revendication 1.

5 15. Protéine de canal sodique codée par un ADN selon les revendications 1 à 9 ou des variants allèles de celui-ci.

16. Protéine de canal sodique résistante à la tétrodotoxine, codée par un ADN selon les revendications 1 à 9 ou des variants allèles de celui-ci.

10 17. Protéine selon la revendication 16, caractérisée en ce qu'elle comprend la séquence d'acides aminés indiquée dans SEQ ID n° 2.

15 18. Procédé d'identification d'inhibiteurs de la protéine de canal sodique résistante à la tétrodotoxine, caractérisé en ce qu'il comprend la mise en contact d'un composé, soupçonné être un tel inhibiteur, avec la protéine de canal sodique de la revendication 16, et la mesure de l'activité de ladite protéine de canal sodique exprimée.

20 19. Anticorps poly- et/ou monoclonaux, suscités contre une protéine de canal sodique résistante à la tétrodotoxine, codée par un ADN des revendications 1 à 9 ou des variants allèles de celui-ci.

25 20. Nécessaire de diagnostic, caractérisé en ce qu'il comprend un polynucléotide selon la revendication 14, capable de s'hybrider spécifiquement avec une protéine de canal sodique résistante à la tétrodotoxine ou un fragment d'une telle protéine.

30 21. Utilisation d'une séquence d'ADN isolée selon les revendications 1 à 9, pour l'identification d'un composé soupçonné être un inhibiteur de protéine de canal sodique résistante à la tétrodotoxine.

1/27

Figure 1A: SEQ ID NO:1

1 GAAGTCACAG GAGTGTCTGT CAGCGAGAGG AAGAAGGGAG AGTTTACTGA
51 GTGTCTTCTG CCCCTCCTCA GGGTGAAGAT GGAGGAGAGG TACTACCCGG
101 TGATCTTCCC GGACGAGCGG AATTTCGGCC CCTTCACTTC CGACTCTCTG
151 GCTGCCATAG AGAAGCGGAT TGCTATCCAA AAGGAGAGGA AGAAGTCCAA
201 AGACAAGGCG GCAGCTGAGC CCCAGCCTCG GCCTCAGCTT GACCTAAAGG
251 CCTCCAGGAA GTTACCTAAG CTTTATGGTG ACATTCCCCC TGAGCTTGTA
301 GCGAAGCCTC TGGGAAGACCT GGACCCATTC TACAAAGACC ATAAGACATT
351 CATGGTGTG AACAGAAGA GAACAATTTA TCGCTTCAGC GCCAAGCGGG
401 CCTTGTTTCAT TCTGGGGCCT TTTAATCCCC TCAGAAGCTT AATGATTTCG
451 ATCTCTGTCC ATTCAGTCTT TAGCATGTTT ATCATCTGCA CGGTGATCAT
501 CAACTGTATG TTCATGGCGA ATTCTATGGA GAGAAGTTTC GACAACGACA
551 TTCCCGAATA CGTCTTCATT GGGATTTATA TTTTAGAAGC TGTGATTAAA
601 ATATTGGCAA GAGGCTTCAT TGTGGATGAG TTTTCCTTCC TCCGAGATCC
651 GTGGAACTGG CTGGACTTCA TTGTCATTGG AACAGCGATC GCAACTTGTT
701 TTCCGGGCAG CCAAGTCAAT CTTTCAGCTC TTCGTACCTT CCGAGTGTTT
751 AGAGCTCTGA AGGCGATTTT AGTTATCTCA GGTCTGAAGG TCATCGTAGG
801 TGCCCTGCTG CGCTCGGTGA AGAAGCTGGT AGACGTGATG GTCCTCACTC
851 TCTTCTGCCT CAGCATCTTT GCCCTGGTCG GTCAGCAGCT GTTCATGGGA
901 ATTCTGAACC AGAAGTGTAT TAAGCACAAC TGTGGCCCCA ACCCTGCATC
951 CAACAAGGAT TGCTTTGAAA AGGAAAAAGA TAGCGAAGAC TTCATAATGT
1001 GTGGTACCTG GCTCGGCAGC AGACCCTGTC CCAATGGTTC TACGTGCGAT
1051 AAAACCACAT TGAACCCAGA CAATAATTAT ACAAAGTTTG ACAACTTTGG
1101 CTGGTCCTTT CTCGCCATGT TCCGGGTTAT GACTCAAGAC TCCTGGGAGA
1151 GGCTTTACCG ACAGATCCTG CGGACCTCTG GGATCTACTT TGTCTTCTTC
1201 TTCGTGGTGG TCATCTTCCT GGGCTCCTTC TACCTGCTTA ACCTAACCCT

2/27

Figure 1B: SEQ ID NO:1

1251 GGCTGTTGTC ACCATGGCTT ATGAAGAACA GAACAGAAAT GTAGCTGCTG
1301 AGACAGAGGC CAAGGAGAAA ATGTTTCAGG AAGCCCAGCA GCTGTTAAGG
1351 GAGGAGAAGG AGGCTCTGGT TGCCATGGGA ATTGACAGAA GTTCCCTTAA
1401 TTCCCTTCAA GCTTCATCCT TTTCCCCGAA GAAGAGGAAG TTTTTCGGTA
1451 GTAAGACAAG AAAGTCCTTC TTTATGAGAG GGTCCAAGAC GGCCCAAGCC
1501 TCAGCGTCTG ATTCAGAGGA CGATGCCTCT AAAAATCCAC AGCTCCTTGA
1551 GCAGACCAAA CGACTGTCCC AGAACTTGCC AGTGGATCTC TTTGATGAGC
1601 ACGTGGACCC CCTCCACAGG CAGAGAGCGC TGAGCGCTGT CAGTATCTTA
1651 ACCATCACCA TGCAGGAACA AGAAAAATTC CAGGAGCCTT GTTTCCCATG
1701 TGGGAAAAAT TTGGCCTCTA AGTACCTGGT GTGGGACTGT AGCCCTCAGT
1751 GGCTGTGCAT AAAGAAGGTC CTGCGGACCA TCATGACGGA TCCCTTTACT
1801 GAGCTGGCCA TCACCATCTG CATCATCATC AATACCGTTT TCTTAGCCGT
1851 GGAGCACCAC AACATGGATG ACAACTTAAA GACCATACTG AAAATAGGAA
1901 ACTGGGTTTT CACGGGAATT TTCATAGCGG AAATGTGTCT CAAGATCATC
1951 GCGCTCGACC CTTACCACTA CTTCCGGCAC GGCTGGAATG TTTTGTACAG
2001 CATCGTGGCC CTCCTGAGTC TCGCTGATGT GCTCTACAAC AACTGTCTG
2051 ATAACAATAG GTCTTTCTTG GCTTCCCTCA GAGTGCTGAG GGTCTTCAAG
2101 TTAGCCAAAT CCTGGCCCAC GTTAAACACT CTCATTAAGA TCATCGGCCA
2151 CTCCGTGGGC GCGCTTGGA ACCTGACTGT GGTCCTGACT ATCGTGGTCT
2201 TCATCTTTTC TGTGGTGGGC ATGCGGCTCT TCGGCACCAA GTTTAACAAG
2251 ACCGCCTACG CCACCCAGGA GCGGCCAGG CGGCGCTGGC ACATGGATAA
2301 TTTCTACCAC TCCTTCCTGG TGGTGTTCGG CATCCTCTGT GGGGAATGGA
2351 TCGAGAACAT GTGGGGCTGC ATGCAGGATA TGGACGGCTC CCCGTTGTGC
2401 ATCATTGTCT TTGTCCTGAT AATGGTGATC GGAAGCTTG TGGTGCTTAA

3/27

Figure 1C: SEQ ID NO:1

2451 CCTCTTCATT GCCTTGCTGC TCAATTCCTT CAGCAATGAG GAGAAGGATG
2501 GGAGCCTGGA AGGAGAGACC AGGAAAACCA AAGTGCAGCT AGCCCTGGAT
2551 CGGTTCCGCC GGGCCTTCTC CTTCATGCTG CACGCTCTTC AGAGTTTTTG
2601 TTGCAAGAAA TGCAGGAGGA AAAACTCGCC AAAGCCAAAA GAGACAACAG
2651 AAAGCTTTGC TGGTGAGAAT AAAGACTCAA TCCTCCCGGA TGCAGGCCCC
2701 TGGAAGGAGT ATGATACAGA CATGGCTTTG TACACTGGAC AGGCCGGGGC
2751 TCCGCTGGCC CCACTCGCAG AGGTAGAGGA CGATGTGGAA TATTGTGGTG
2801 AAGGCGGTGC CCTACCCACC TCACAACATA GTGCTGGAGT TCAGGCCGGT
2851 GACCTCCCTC CAGAGACCAA GCAGCTCACT AGCCCGGATG ACCAAGGGGT
2901 TGAAATGGAA GTATTTTCTG AAGAAGATCT GCATTTAAGC ATACAGAGTC
2951 CTCGAAAGAA GTCTGACGCA GTGAGCATGC TCTCGGAATG CAGCACAATT
3001 GACCTGAATG ATATCTTTAG AAATTTACAG AAAACAGTTT CCCCCAAAAA
3051 GCAGCCAGAT AGATGCTTTC CCAAGGGCCT TAGTTGTCAC TTTCTATGCC
3101 ACAAACAGA CAAGAGAAAG TCCCCCTGGG TCCTGTGGTG GAACATTCGG
3151 AAAACCTGCT ACCAAATCGT GAAGCACAGC TGGTTTGAGA GTTTCATAAT
3201 CTTTGTTATT CTGCTGAGCA GTGGAGCGCT GATATTTGAA GATGTCAATC
3251 TCCCCAGCCG GCCCCAAGTT GAGAAATTAC TAAGGTGTAC CGATAATATT
3301 TTCACATTTA TTTTCCTCCT GGAAATGATC CTGAAGTGGG TGGCCTTTGG
3351 ATTCCGGAGG TATTTACCA GTGCCTGGTG CTGGCTTGAT TTCCTCATTG
3401 TGGTGGTGTC TGTGCTCAGT CTCATGAATC TACCAAGCTT GAAGTCCTTC
3451 CGGACTCTGC GGGCCCTGAG ACCTCTGCGG GCGCTGTCCC AGTTTGAAGG
3501 AATGAAGGTT GTCGTCTACG CCCTGATCAG CGCCATACCT GCCATTCTCA
3551 ATGTCTTGCT GGTCTGCCTC ATTTTCTGGC TCGTATTTTG TATCTTGGGA
3601 GTAAATTTAT TTTCTGGGAA GTTTGGAAGG TGCATTAACG GGACAGACAT

Figure 1D: SEQ ID NO:1

3651 AAATATGTAT TTGGATTTTA CCGAAGTTCC GAACCGAAGC CAATGTAACA
3701 TTAGTAATTA CTCGTGGAAG GTCCCGCAGG TCAACTTTGA CAACGTGGGG
3751 AATGCCTATC TCGCCCTGCT GCAAGTGGCA ACCTATAAGG GCTGGCTGGA
3801 AATCATGAAT GCTGCTGTCG ATTCCAGAGA GAAAGACGAG CAGCCGGACT
3851 TTGAGGCGAA CCTCTACGCG TATCTCTACT TTGTGGTTTT TATCATCTTC
3901 GGCTCCTTCT TTACCCTGAA CCTCTTTATC GGTGTTATTA TTGACAACCT
3951 CAATCAGCAG CAGAAAAAGT TAGGTGGCCA AGACATTTTT ATGACAGAAG
4001 AACAGAAGAA ATATTACAAT GCAATGAAAA AGTTAGGAAC CAAGAAACCT
4051 CAAAAGCCCA TCCCAAGGCC CCTGAACAAA TGTCAAGCCT TTGTGTTCGA
4101 CCTGGTCACA AGCCAGGTCT TTGACGTCAT CATTCTGGGT CTTATTGTCT
4151 TAAATATGAT TATCATGATG GCTGAATCTG CCGACCAGCC CAAAGATGTG
4201 AAGAAAACCT TTGATATCCT CAACATAGCC TTCGTGGTCA TCTTTACCAT
4251 AGAGTGTCTC ATCAAAGTCT TTGCTTTGAG GCAACACTAC TTCACCAATG
4301 GCTGGAACCT ATTTGATTGT GTGGTCGTGG TTCTTTCTAT CATTAGTACC
4351 CTGGTTTCCC GCTTGGAGGA CAGTGACATT TCTTTCCCGC CCACGCTCTT
4401 CAGAGTCGTC CGCTTGGCTC GGATTGGTCG AATCCTCAGG CTGGTCCGGG
4451 CTGCCCCGGG AATCAGGACC CTCCTCTTTG CTTTGATGAT GTCTCTCCCC
4501 TCTCTCTTCA ACATCGGTCT GCTGCTCTTC CTGGTGATGT TCATTTACGC
4551 CATCTTTGGG ATGAGCTGGT TTTCCAAAGT GAAGAAGGGC TCCGGGATCG
4601 ACGACATCTT CAACTTCGAG ACCTTTACGG GCAGCATGCT GTGCCTCTTC
4651 CAGATAACCA CTTCGGCTGG CTGGGATACC CTCCTCAACC CCATGCTGGA
4701 GGCAAAAGAA CACTGCAACT CCTCCTCCCA AGACAGCTGT CAGCAGCCGC
4751 AGATAGCCGT CGTCTACTTC GTCAGTTACA TCATCATCTC CTCCTCATC
4801 GTGGTCAACA TGTACATCGC TGTGATCCTC GAGAACTTCA ACACAGCCAC

5/27

Figure 1E: SEQ ID NO: 1

4851 GGAGGAGAGC GAGGACCCTC TGGGAGAGGA CGACTTTGAA ATCTTCTATG
4901 AGGTCTGGGA GAAGTTTGAC CCCGAGGCGT CGCAGTTCAT CCAGTATTCG
4951 GCCCTCTCTG ACTTTGCGGA CGCCCTGCCG GAGCCGTTGC GTGTGGCCAA
5001 GCCGAATAAG TTTCAGTTTC TAGTGATGGA CTTGCCCATG GTGATGGGCG
5051 ACCGCCTCCA TTGCATGGAT GTTCTCTTTG CTTTCACTAC CAGGGTCCTC
5101 GGGGACTCCA GCGGCTTGGA TACCATGAAA ACCATGATGG AGGAGAAGTT
5151 TATGGAGGCC AACCCTTTTA AGAAGCTCTA CGAGCCCATA GTCACCACCA
5201 CCAAGAGGAA GGAGGAGGAG CAAGGCGCCG CCGTCATCCA GAGGGCCTAC
5251 CGGAAACACA TGGAGAAGAT GGTCAAACTG AGGCTGAAGG ACAGGTCAAG
5301 TTCATCGCAC CAGGTGTTTT GCAATGGAGA CTTGTCCAGC TTGGATGTGG
5351 CCAAGGTCAA GGTTACAAT GACTGAACCC TCATCTCCAC CCCTACCTCA
5401 CTGCCTCACA GCTTAGCCTC CAGCCTCTGG CGAGCAGGCG GCAGACTCAC
5451 TGAACACAGG CCGTTCGATC TGTGTTTTTG GCTGAACGAG GTGACAGGTT
5501 GCGGTCCATT TTAAATGAC TCTTGAAAG ATTTTCATGTA GAGAGATGTT
5551 AGAAGGGACT GCAAAGGACA CCGACCATAA CGGAAGGCCT GGAGGACAGT
5601 CCAACTTACA TAAAGATGAG AAACAAGAAG GAAAGATCCC AGGAAAACCTT
5651 CAGATTGTGT TCTCAGTACA TCCCCAATG TGTCTGTTTC GTGTTTTGAG
5701 TATGTGACCT GCCACATGTA GCTCTTTTTT GCATGTACGT CAAAACCCTG
5751 CAGTAAGTTG ATAGCTTGCT ACGGGTGTTT CTACCAGCAT CACAGAATTG
5801 GGTGTATGAC TCAAACCTAA AAGCATGACT CTGACTTGTC AGTCAGCACC
5851 CCGACTTTCA GACGCTCCAA TCTCTGTCCC AGGTGTCTAA CGAATAAATA
5901 GGTAAG

Met	Glu	Glu	Arg	Tyr	Tyr	Pro	Val	Ile	Phe	Pro	Asp	Glu	Arg	Asn	Phe
1				5					10					15	
Arg	Pro	Phe	Thr	Ser	Asp	Ser	Leu	Ala	Ala	Ile	Glu	Lys	Arg	Ile	Ala
			20					25					30		
Ile	Gln	Lys	Glu	Arg	Lys	Lys	Ser	Lys	Asp	Lys	Ala	Ala	Ala	Glu	Pro
		35					40				45				
Gln	Pro	Arg	Pro	Gln	Leu	Asp	Leu	Lys	Ala	Ser	Arg	Lys	Leu	Pro	Lys
	50					55					60				
Leu	Tyr	Gly	Asp	Ile	Pro	Pro	Glu	Leu	Val	Ala	Lys	Pro	Leu	Glu	Asp
65				70						75				80	
Leu	Asp	Pro	Phe	Tyr	Lys	Asp	His	Lys	Thr	Phe	Met	Val	Leu	Asn	Lys
			85						90					95	
Lys	Arg	Thr	Ile	Tyr	Arg	Phe	Ser	Ala	Lys	Arg	Ala	Leu	Phe	Ile	Leu
		100						105					110		
Gly	Pro	Phe	Asn	Pro	Leu	Arg	Ser	Leu	Met	Ile	Arg	Ile	Ser	Val	His
	115						120					125			
Ser	Val	Phe	Ser	Met	Phe	Ile	Ile	Cys	Thr	Val	Ile	Ile	Asn	Cys	Met
	130					135					140				
Phe	Met	Ala	Asn	Ser	Met	Glu	Arg	Ser	Phe	Asp	Asn	Asp	Ile	Pro	Glu
145				150						155				160	
Tyr	Val	Phe	Ile	Gly	Ile	Tyr	Ile	Leu	Glu	Ala	Val	Ile	Lys	Ile	Leu
			165						170					175	
Ala	Arg	Gly	Phe	Ile	Val	Asp	Glu	Phe	Ser	Phe	Leu	Arg	Asp	Pro	Trp
		180						185					190		
Asn	Trp	Leu	Asp	Phe	Ile	Val	Ile	Gly	Thr	Ala	Ile	Ala	Thr	Cys	Phe
	195						200					205			
Pro	Gly	Ser	Gln	Val	Asn	Leu	Ser	Ala	Leu	Arg	Thr	Phe	Arg	Val	Phe
	210					215					220				
Arg	Ala	Leu	Lys	Ala	Ile	Ser	Val	Ile	Ser	Gly	Leu	Lys	Val	Ile	Val
225				230						235				240	
Gly	Ala	Leu	Leu	Arg	Ser	Val	Lys	Lys	Leu	Val	Asp	Val	Met	Val	Leu
			245						250					255	
Thr	Leu	Phe	Cys	Leu	Ser	Ile	Phe	Ala	Leu	Val	Gly	Gln	Gln	Leu	Phe
		260						265					270		
Met	Gly	Ile	Leu	Asn	Gln	Lys	Cys	Ile	Lys	His	Asn	Cys	Gly	Pro	Asn
	275						280					285			

2771103

7/27

Pro Ala .Ser Asn Lys Asp Cys Phe Glu Lys Glu Lys Asp Ser Glu Asp

290

295

300

Phe Ile Met Cys Gly Thr Trp Leu Gly Ser Arg Pro Cys Pro Asn Gly

305

310

315

320

8/27

Figure 2B: SEQ ID NO: 2

```

Ser Thr Cys Asp Lys Thr Thr Leu Asn Pro Asp Asn Asn Tyr Thr Lys
      325                      330                      335
Phe Asp Asn Phe Gly Trp Ser Phe Leu Ala Met Phe Arg Val Met Thr
      340                      345                      350
Gln Asp Ser Trp Glu Arg Leu Tyr Arg Gln Ile Leu Arg Thr Ser Gly
      355                      360                      365
Ile Tyr Phe Val Phe Phe Phe Val Val Val Ile Phe Leu Gly Ser Phe
      370                      375                      380
Tyr Leu Leu Asn Leu Thr Leu Ala Val Val Thr Met Ala Tyr Glu Glu
      385                      390                      395                      400
Gln Asn Arg Asn Val Ala Ala Glu Thr Glu Ala Lys Glu Lys Met Phe
      405                      410                      415
Gln Glu Ala Gln Gln Leu Leu Arg Glu Glu Lys Glu Ala Leu Val Ala
      420                      425                      430
Met Gly Ile Asp Arg Ser Ser Leu Asn Ser Leu Gln Ala Ser Ser Phe
      435                      440                      445
Ser Pro Lys Lys Arg Lys Phe Phe Gly Ser Lys Thr Arg Lys Ser Phe
      450                      455                      460
Phe Met Arg Gly Ser Lys Thr Ala Gln Ala Ser Ala Ser Asp Ser Glu
      465                      470                      475                      480
Asp Asp Ala Ser Lys Asn Pro Gln Leu Leu Glu Gln Thr Lys Arg Leu
      485                      490                      495
Ser Gln Asn Leu Pro Val Asp Leu Phe Asp Glu His Val Asp Pro Leu
      500                      505                      510
His Arg Gln Arg Ala Leu Ser Ala Val Ser Ile Leu Thr Ile Thr Met
      515                      520                      525
Gln Glu Gln Glu Lys Phe Gln Glu Pro Cys Phe Pro Cys Gly Lys Asn
      530                      535                      540
Leu Ala Ser Lys Tyr Leu Val Trp Asp Cys Ser Pro Gln Trp Leu Cys
      545                      550                      555                      560
Ile Lys Lys Val Leu Arg Thr Ile Met Thr Asp Pro Phe Thr Glu Leu
      565                      570                      575
Ala Ile Thr Ile Cys Ile Ile Ile Asn Thr Val Phe Leu Ala Val Glu
      580                      585                      590
His His Asn Met Asp Asp Asn Leu Lys Thr Ile Leu Lys Ile Gly Asn

```


2771103

9/27

595

600

605

Trp Val Phe Thr Gly Ile Phe Ile Ala Glu Met Cys Leu Lys Ile Ile

610

615

620

10/27

Figure 2C: SEQ ID NO: 2

Ala Leu Asp Pro Tyr His Tyr Phe Arg His Gly Trp Asn Val Phe Asp
 625 630 635 640
 Ser Ile Val Ala Leu Leu Ser Leu Ala Asp Val Leu Tyr Asn Thr Leu
 645 650 655
 Ser Asp Asn Asn Arg Ser Phe Leu Ala Ser Leu Arg Val Leu Arg Val
 660 665 670
 Phe Lys Leu Ala Lys Ser Trp Pro Thr Leu Asn Thr Leu Ile Lys Ile
 675 680 685
 Ile Gly His Ser Val Gly Ala Leu Gly Asn Leu Thr Val Val Leu Thr
 690 695 700
 Ile Val Val Phe Ile Phe Ser Val Val Gly Met Arg Leu Phe Gly Thr
 705 710 715 720
 Lys Phe Asn Lys Thr Ala Tyr Ala Thr Gln Glu Arg Pro Arg Arg Arg
 725 730 735
 Trp His Met Asp Asn Phe Tyr His Ser Phe Leu Val Val Phe Arg Ile
 740 745 750
 Leu Cys Gly Glu Trp Ile Glu Asn Met Trp Gly Cys Met Gln Asp Met
 755 760 765
 Asp Gly Ser Pro Leu Cys Ile Ile Val Phe Val Leu Ile Met Val Ile
 770 775 780
 Gly Lys Leu Val Val Leu Asn Leu Phe Ile Ala Leu Leu Leu Asn Ser
 785 790 795 800
 Phe Ser Asn Glu Glu Lys Asp Gly Ser Leu Glu Gly Glu Thr Arg Lys
 805 810 815
 Thr Lys Val Gln Leu Ala Leu Asp Arg Phe Arg Arg Ala Phe Ser Phe
 820 825 830
 Met Leu His Ala Leu Gln Ser Phe Cys Cys Lys Lys Cys Arg Arg Lys
 835 840 845
 Asn Ser Pro Lys Pro Lys Glu Thr Thr Glu Ser Phe Ala Gly Glu Asn
 850 855 860
 Lys Asp Ser Ile Leu Pro Asp Ala Arg Pro Trp Lys Glu Tyr Asp Thr
 865 870 875 880
 Asp Met Ala Leu Tyr Thr Gly Gln Ala Gly Ala Pro Leu Ala Pro Leu
 885 890 895

2771103

11/27

Ala Glu Val Glu Asp Asp Val Glu Tyr Cys Gly Glu Gly Gly Ala Leu

900

905

910

12/27
Figure 2D: SEQ ID NO: 2

Pro Thr Ser Gln His Ser Ala Gly Val Gln Ala Gly Asp Leu Pro Pro
 915 920 925
 Glu Thr Lys Gln Leu Thr Ser Pro Asp Asp Gln Gly Val Glu Met Glu
 930 935 940
 Val Phe Ser Glu Glu Asp Leu His Leu Ser Ile Gln Ser Pro Arg Lys
 945 950 955 960
 Lys Ser Asp Ala Val Ser Met Leu Ser Glu Cys Ser Thr Ile Asp Leu
 965 970 975
 Asn Asp Ile Phe Arg Asn Leu Gln Lys Thr Val Ser Pro Lys Lys Gln
 980 985 990
 Pro Asp Arg Cys Phe Pro Lys Gly Leu Ser Cys His Phe Leu Cys His
 995 1000 1005
 Lys Thr Asp Lys Arg Lys Ser Pro Trp Val Leu Trp Trp Asn Ile Arg
 1010 1015 1020
 Lys Thr Cys Tyr Gln Ile Val Lys His Ser Trp Phe Glu Ser Phe Ile
 1025 1030 1035 1040
 Ile Phe Val Ile Leu Leu Ser Ser Gly Ala Leu Ile Phe Glu Asp Val
 1045 1050 1055
 Asn Leu Pro Ser Arg Pro Gln Val Glu Lys Leu Leu Arg Cys Thr Asp
 1060 1065 1070
 Asn Ile Phe Thr Phe Ile Phe Leu Leu Glu Met Ile Leu Lys Trp Val
 1075 1080 1085
 Ala Phe Gly Phe Arg Arg Tyr Phe Thr Ser Ala Trp Cys Trp Leu Asp
 1090 1095 1100
 Phe Leu Ile Val Val Val Ser Val Leu Ser Leu Met Asn Leu Pro Ser
 1105 1110 1115 1120
 Leu Lys Ser Phe Arg Thr Leu Arg Ala Leu Arg Pro Leu Arg Ala Leu
 1125 1130 1135
 Ser Gln Phe Glu Gly Met Lys Val Val Val Tyr Ala Leu Ile Ser Ala
 1140 1145 1150
 Ile Pro Ala Ile Leu Asn Val Leu Leu Val Cys Leu Ile Phe Trp Leu
 1155 1160 1165
 Val Phe Cys Ile Leu Gly Val Asn Leu Phe Ser Gly Lys Phe Gly Arg
 1170 1175 1180
 Cys Ile Asn Gly Thr Asp Ile Asn Met Tyr Leu Asp Phe Thr Glu Val

2771103

13/27

1185

1190

1195

1200

Pro Asn Arg Ser Gln Cys Asn Ile Ser Asn Tyr Ser Trp Lys Val Pro

1205

1210

1215

14/27
Figure 2E: SEQ ID NO: 2

Gln Val Asn Phe Asp Asn Val Gly Asn Ala Tyr Leu Ala Leu Leu Gln
1220 1225 1230

Val Ala Thr Tyr Lys Gly Trp Leu Glu Ile Met Asn Ala Ala Val Asp
1235 1240 1245

Ser Arg Glu Lys Asp Glu Gln Pro Asp Phe Glu Ala Asn Leu Tyr Ala
1250 1255 1260

Tyr Leu Tyr Phe Val Val Phe Ile Ile Phe Gly Ser Phe Phe Thr Leu
1265 1270 1275 1280

Asn Leu Phe Ile Gly Val Ile Ile Asp Asn Phe Asn Gln Gln Gln Lys
1285 1290 1295

Lys Leu Gly Gly Gln Asp Ile Phe Met Thr Glu Glu Gln Lys Lys Tyr
1300 1305 1310

Tyr Asn Ala Met Lys Lys Leu Gly Thr Lys Lys Pro Gln Lys Pro Ile
1315 1320 1325

Pro Arg Pro Leu Asn Lys Cys Gln Ala Phe Val Phe Asp Leu Val Thr
1330 1335 1340

Ser Gln Val Phe Asp Val Ile Ile Leu Gly Leu Ile Val Leu Asn Met
1345 1350 1355 1360

Ile Ile Met Met Ala Glu Ser Ala Asp Gln Pro Lys Asp Val Lys Lys
1365 1370 1375

Thr Phe Asp Ile Leu Asn Ile Ala Phe Val Val Ile Phe Thr Ile Glu
1380 1385 1390

Cys Leu Ile Lys Val Phe Ala Leu Arg Gln His Tyr Phe Thr Asn Gly
1395 1400 1405

Trp Asn Leu Phe Asp Cys Val Val Val Val Leu Ser Ile Ile Ser Thr
1410 1415 1420

Leu Val Ser Arg Leu Glu Asp Ser Asp Ile Ser Phe Pro Pro Thr Leu
1425 1430 1435 1440

Phe Arg Val Val Arg Leu Ala Arg Ile Gly Arg Ile Leu Arg Leu Val
1445 1450 1455

Arg Ala Ala Arg Gly Ile Arg Thr Leu Leu Phe Ala Leu Met Met Ser
1460 1465 1470

Leu Pro Ser Leu Phe Asn Ile Gly Leu Leu Leu Phe Leu Val Met Phe
1475 1480 1485

Ile Tyr Ala Ile Phe Gly Met Ser Trp Phe Ser Lys Val Lys Lys Gly

2771103

15/27

1490

1495

1500

Ser Gly Ile Asp Asp Ile Phe Asn Phe Glu Thr Phe Thr Gly Ser Met

1505

- 1510

1515

1520

16/27
Figure 2F: SEQ ID NO: 2

Leu Cys Leu Phe Gln Ile Thr Thr Ser Ala Gly Trp Asp Thr Leu Leu
 1525 1530 1535
 Asn Pro Met Leu Glu Ala Lys Glu His Cys Asn Ser Ser Ser Gln Asp
 1540 1545 1550
 Ser Cys Gln Gln Pro Gln Ile Ala Val Val Tyr Phe Val Ser Tyr Ile
 1555 1560 1565
 Ile Ile Ser Phe Leu Ile Val Val Asn Met Tyr Ile Ala Val Ile Leu
 1570 1575 1580
 Glu Asn Phe Asn Thr Ala Thr Glu Glu Ser Glu Asp Pro Leu Gly Glu
 1585 1590 1595 1600
 Asp Asp Phe Glu Ile Phe Tyr Glu Val Trp Glu Lys Phe Asp Pro Glu
 1605 1610 1615
 Ala Ser Gln Phe Ile Gln Tyr Ser Ala Leu Ser Asp Phe Ala Asp Ala
 1620 1625 1630
 Leu Pro Glu Pro Leu Arg Val Ala Lys Pro Asn Lys Phe Gln Phe Leu
 1635 1640 1645
 Val Met Asp Leu Pro Met Val Met Gly Asp Arg Leu His Cys Met Asp
 1650 1655 1660
 Val Leu Phe Ala Phe Thr Thr Arg Val Leu Gly Asp Ser Ser Gly Leu
 1665 1670 1675 1680
 Asp Thr Met Lys Thr Met Met Glu Glu Lys Phe Met Glu Ala Asn Pro
 1685 1690 1695
 Phe Lys Lys Leu Tyr Glu Pro Ile Val Thr Thr Thr Lys Arg Lys Glu
 1700 1705 1710
 Glu Glu Gln Gly Ala Ala Val Ile Gln Arg Ala Tyr Arg Lys His Met
 1715 1720 1725
 Glu Lys Met Val Lys Leu Arg Leu Lys Asp Arg Ser Ser Ser Ser His
 1730 1735 1740
 Gln Val Phe Cys Asn Gly Asp Leu Ser Ser Leu Asp Val Ala Lys Val
 1745 1750 1755 1760
 Lys Val His Asn Asp
 1765

17/27

Figure 2G: SEQ ID NO:2

```

1  MEERYYPVIF PDERNFRPFT SDSLAAIEKR IAIQKERKKS KDKAAAEQPF
51  RPQLDLKASR KLPKLYGDIP PELVAKPLED LDPFYKDHKT FMVLNKKRTI
101 YRFSAKRALF ILGPFNPLRS LMIRISVHSV FSMFIICTVI INCMFMANSM
    |-----IS1-----|
151 ERSFDNDIPE YVFIGIYILE AVIKILARGF IVDEFSFLRD PWNWLDIFI
    |-----IS2-----| |-----IS3
201 GTAIATCFPG SQVNLSALRT FRVFRALKAI SVISGLKVIV GALLRSVKKL
    -----| • |-----IS4-----|
251 VDMVMTLFC LSIFALVGQQ LFMGILNQKC IKHNCGPNPA SNKDCFEKEK
    |-----IS5-----|
301 DSEDFIMCGT WLGSRPCPNG STCDKTTLNP DNNYTKFDNF GWSFLAMFRV
    •
351 MTQDSWERLY RQILRTSGIY FVFFVVFVIF LGSFYLLNLT LAVVTMAYEE
    |-----IS6-----|
401 QNRNVAAETE AKEKMFQEAQ QLLREEKEAL VAMGIDRSSL NSLQASSFSP
451 KKRKFFGSKT RKSFFMRGSK TAQASASDSE DDASKNPQLL EQTKRLSQNL
    •
501 PVDLFDEHVD PLHRQRALSA VSILTITMQE QEKFQEPCCP CGKNLASKYL
551 VWDCSPQWLC IKKVLRTIMT DPFTELAITY CIIINTVFLA VEHNMDDNL
    |-----IIS1-----|
601 KTIKIGNWV FTGIFIAEMC LKIIALDPYH YFRHGWNVFD SIVALLSLAD
    |-----IIS2-----| |-----IIS3-
651 VLYNTLSDMN RSFLASLRVL RVFKLAKSWP TLNTLIKIIG HSVGALGNLT
    ----| • |-----IIS4-----| •|
701 VVLTIVVFIF SVVGMRLFGT KFNKTAYATQ ERPRRWHTD NFYHSFLVVF
    -----IIS5-----| •
751 RILCGEWIEN MWGCMQDMG SPLCIIVFVL IMVIGKLVVL NLFIALLLNS
    |-----IIS6-----|
801 FSNEEKDGS L EGETRKTQVQ LALDRFRRF SFMLHALQSF CCKKCRRKNS
    •
851 PKPKETTESF AGENKDSILP DARPWKEYDT DMALYTGQAG APLAPLAEVE
901 DDVEYCGEGG ALPTSQHSAG VQAGDLPPET KQLTSPDDQG VEMEVSSEED
951 LHLSIQSPRK KSDAVSMLSE CSTIDLNDIF RNLQKTVSPK KQPDRCFPKG
    •
1001 LSCHFLCHKT DKRKSPWVLW WNIRKTCYQI VKHSWFESFI IFVILLSSGA
    |-----IIS1-----|
1051 LIFEDVNLPS RPQVEKLLRC TDNIFTFIFL LEMILKWVAF GFRRYFTSAW
    --| |-----IIS2-----| |
1101 CWLDFLIWV SVLSLMNLPS LKSFRTLRL RPLRLSQFE GMKVVVYALI
    ----IIS3-----| |-----IIS4-----|
1151 SAIPAILNVL LVCLIFWLVF CILGVNLFSG KFGRCINGTD INMYLDFTEV
    |-----IIS5-----| •
1201 PNRSQCNISN YSWKVPQVNF DNVGNAYLAL LQVATYKWL EIMNAAVDSR
    • • •
1251 EKDEQPDFFA NLYAYLYFVV FIIFGSFFTL NLFIVGIIDN FNQQQKKLGG
    |-----IIS6-----|

```

18/27
Figure 2H: SEQ ID NO: 2

```
1301 QDIFMTEEQK KYYNAMKKLG TKKPQKPIPR PLNKCQAFVF DLVTSQVFDV
                                           |-----
1351 IILGLIVLNM IIMMAESADQ PKDVKKTFDI LNIAFVVIFT IECLIKVFAL
IVS1-----| |-----IVS2-----|
1401 RQHYFTNGWN LFDCVVVVL IISTLVSRL DSDISFPPTL FRVVRLARIG
|-----IVS3-----| |-----
1451 RILRLVRAAR GIRTLLFALM MSLPSLFNIG LLLFLVMFIY AIFGMSWFSK
IVS4-----| |-----IVS5-----
1501 VKKGS GIDDI FNFETFTGSM LCLFQITTS GWDTLNPMLEAKEHCNSSS
| O
1551 QDSCQQPQIA VVYFVSYIII SFLIVVNYI AVILENFNTA TEESEDPLGE
|-----IVS6-----|
1601 DDFEIFYEVW EKFDPEASQF IQYSALS DFDALPEPLRVA KPNKFQFLVM
1651 DLPMVMGDRL HCMDVLFAFT TRVLGDSSGL DTMKTMMEEK FMEANPFKKL
1701 YEPIVTTTKR KEEEQGA AVI QRAYRKHMEK MVKLSLKDRS SSSHQVFCNG
1751 DLSSLDVAKV KVHND*
```

19/27

Figure 3A: SEQ ID NO:3

1 GCTGAGCAGT GGGGCACTGA TATTTGAAGA TGTTCACCTT GAGAACCAAC
51 CCAAAATCCA AGAATTACTA AATTGTA CTG ACATTATTTT TACACATATT
101 TTTATCCTGG AGATGGTACT AAAATGGGTA GCCTTCGGAT TTGGAAAGTA
151 TTTCAACCAGT GCCTGGTGCT GCCTTGATTT CATCATTTGTG ATTGTCTCTG
201 TGACCACCCT CATTAACCTA ATGGAATTGA AGTCCTTCCG GACTCTACGA
251 GCACTGAGGC CTCTTCGTGC GCTGTCCCAG TTTGAAGGAA TGAAGGTGGT
301 GGTCAATGCT CTCATAGGTG CCATACCTGC CATTCTGAAT GTTTTGCTTG
351 TCTGCCTCAT TTTCTGGCTC GTATTTTGTA TTCTGGGAGT ATACTTCTTT
401 TCTGGAAAAT TTGGGAAATG CATTAATGGA ACAGACTCAG TTATAAATTA
451 TACCATCATT ACAAATAAAA GTCAATGTGA AAGTGGCAAT TTCTCTTGGA
501 TCAACCAGAA AGTCAACTTT GACAATGTGG GAAATGCTTA CCTCGCTCTG
551 CTGCAAGTGG CAACATTTAA GGGCTGGATG GATATTATAT ATGCAGCTGT
601 TGATTCCACA GAGAAAGAAC AACAGCCAGA GTTTGAGAGC AATTCACCTG
651 GTTACATTTA CTTCGTAGTC TTTATCATCT TTGGCTCATT CTTCACCTCTG
701 AATCTCTTCA TTGGCGTTAT CATTGACAAC TTCAACCAAC AGCAGAAAAA
751 GTTAGGTGGC CAAGACATTT TTATGACAGA AGAACAGAAG AAATACTATA
801 ATGCAATGAA AAAATTAGGA TCCAAAAAAC CTCAAAAACC CATTCCACGG
851 CCCGTT

(Ligne supérieure: PN5 humain)
(Ligne inférieure: PN5 de rat)

```

1 .....LSSGA 5
    |||||
1001 LSCHFLCHKTDKRKSPWVLWNNIRKTCYQIVKHSWFESFIIFVILLSSGA 1050
    .
    6 LIFEDVHLENQPKIQELLNCTDIIFTHIFILEMVLKWVAFGFGKYFTSAW 55
    |||||. | .|. :. || ||| ||| ||: ||: ||||| | : |||||
1051 LIFEDVNLPSRPQVEKLLRCTDNIFTFIFLLEMILKWVAFGFRRYFTSAW 1100
    .
    56 CCLDFIIVIVSVTTLINLMELKSFRTLRLRPLRLALSQFEGMKVVVNALI 105
    | |||: ||: ||| .|. || ||||| ||||| ||||| ||||| |||||
1101 CWLDFLIVVVSVLSLMNLPSLKSFRTLRLRPLRLALSQFEGMKVVVYALI 1150
    .
    106 GAIPAILNVLLVCLIFWLVCILGVYFFSGKFGKCINGTD..SVINYTII 153
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| :. : |
1151 SAIPAILNVLLVCLIFWLVCILGVNLFSGKFGRCINGTDINMYLDFTEV 1200
    .
    154 TNKSQCESGNFSWINQKVNFNDVGNAYLALLQVATFKGWMDIIYA AVDST 203
    |: ||| |: || . ||||| ||||| ||||| ||||| : ||| :. |||||
1201 PNRSQCNISNYSWKVPQVNFNDVGNAYLALLQVATYKGWLEIMNA AVDSR 1250
    .
    204 EKEQQPEFESNSLGYIYFVVFIIFGSFFTLNLFIGVIIDNFNQQQKKLGG 253
    ||: ||: ||. | |: ||||| ||||| ||||| ||||| ||||| |||||
1251 EKDEQPDFEANLYAYLYFVVFIIFGSFFTLNLFIGVIIDNFNQQQKKLGG 1300
    .
    254 QDIFMTEEQKKYYNAMKKLGSKKPQKPIPRPV..... 285
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1301 QDIFMTEEQKKYYNAMKKLGTKKPQKPIPRPLNKCQAFVFDLVT SQVFDV 1350

```

Figure 4: SEQ ID NO:4

```
1   CTCAACATGG TTACGATGAT GGTGGAGACC GACGAGCAGG GCGAGGAGAA
51  GACGAAGGTT CTGGGCAGAA TCAACCAGTT CTTTGTGGCC GTCTTCACGG
101 GCGAGTGTGT GATGAAGATG TTCGCCCTGC GACAGTACTA TTTACCAAC
151 GGCTGGAACG TGTTCGAcTT CATAGTGGTG ATCCTGTCCA TTGGGAGTCT
201 GCTGTTTCT  GCAATCCTTA AGTCACTGGA AAACACTTC TCCCCGACGC
251 TCTTCCGGGT CATCCGTCTG GCCAGGATCG GCCGCATCCT CAGGCTGATC
301 CGAGCAGCCA AGGGGATTCTG CACGCTGCTC TTCGCCCTCA TGATGTCCCT
351 GCCCCGCCCTC TTCAACATCG GCCTCCTCCT CTTCTCTGtC ATGTTCATCT
401 ACTCCATCTT CGGCATGGCC AGCTTCGCTA ACGTCGTGGA CGAGGCCGGC
451 ATCGACGACA TGTTCAACTT CAAGACCTTT GGCAACAGCA TGCTGTGCCT
501 GTTCCAGATC ACCACCTCGG CCGGCTGGGA CGGCCTCCTC AGCCCCATCC
551 TCAACACGGG GCCTCCCTAC TGCGACCCCA ACCTGCCCAA CAGCAACGGC
601 TCCCGGGGGA ACTGCGGGAG CCCGGCGGTG GGCATCATCT TCTTCACCAC
651 CTACATCATC ATCTCCTTCC TCATCGTGGT CAACATGTAT ATCGCAGTCA
701 TC
```

Figure 5A: SEQ ID NO: 5

1 GTCGACTCTA GATCAGGGTG AAGATGGAGG AGAGGTACTA CCCGGTGATC
51 TTCCCGGACG AGCGGAATTT CCGCCCCTTC ACTTCCGACT CTCTGGCTGC
101 CATAGAGAAG CGGATTGCTA TCCAAAAGGA GAGGAAGAAG TCCAAAGACA
151 AGGCGGCAGC TGAGCCCCAG CCTCGGCCTC AGCTTGACCT AAAGGCCTCC
201 AGGAAGTTAC CTAAGCTTTA TGGTGACATT CCCCCTGAGC TTGTAGCGAA
251 GCCTCTGGAA GACCTGGACC CATTCTACAA AGACCATAAG ACATTCATGG
301 TGTTGAACAA GAAGAGAACA ATTTATCGCT TCAGCGCCAA GCGGGCCTTG
351 TTCATTCTGG GGCCTTTTAA TCCCCTCAGA AGCTTAATGA TTCGTATCTC
401 TGTCCATTCA GTCTTTAGCA TGTTTCATCAT CTGCACGGTG ATCATCAACT
451 GTATGTTTCAT GGCGAATTCT ATGGAGAGAA GTTTCGACAA CGACATTCCC
501 GAATACGTCT TCATTGGGAT TTATATTTTA GAAGCTGTGA TTAATAATATT
551 GGCAAGAGGC TTCATTGTGG ATGAGTTTTT CTTCTCCGA GATCCGTGGA
601 ACTGGCTGGA CTTCAATTGTC ATTGGAACAG CGATCGCAAC TTGTTTTCCG
651 GGCAGCCAAG TCAATCTTTC AGCTCTTCGT ACCTTCCGAG TGTTCAGAGC
701 TCTGAAGGCG ATTTTCAGTTA TCTCAGGTCT GAAGGTCATC GTAGGTGCCC
751 TGCTGCGCTC GGTGAAGAAG CTGGTAGACG TGATGGTCCT CACTCTCTTC
801 TGCCTCAGCA TCTTTGCCCT GGTGCGTCAG CAGCTGTTCA TGGGAATTCT
851 GAACCAGAAG TGTATTAAGC ACAACTGTGG CCCCACCCT GCATCCAACA
901 AGGATTGCTT TGAAAAGGAA AAAGATAGCG AAGACTTCAT AATGTGTGGT
951 ACCTGGCTCG GCAGCAGACC CTGTCCCAAT GGTCTACGT GCGATAAAAC
1001 CACATTGAAC CCAGACAATA ATTATACAAA GTTGACAAC TTTGGCTGGT
1051 CCTTTCTCGC CATGTTCCGG GTTATGACTC AAGACTCCTG GGAGAGGCTT
1101 TACCGACAGA TCCTGCGGAC CTCTGGGATC TACTTTGTCT TCTTCTTCGT

23/27
Figure 5B: SEQ ID NO: 5

1151 GGTGGTCATC TTCCTGGGCT CCTTCTACCT GCTTAACCTA ACCCTGGCTG
1201 TTGTCACCAT GGCTTATGAA GAACAGAACA GAAATGTAGC TGCTGAGACA
1251 GAGGCCAAGG AGAAAATGTT TCAGGAAGCC CAGCAGCTGT TAAGGGAGGA
1301 GAAGGAGGCT CTGGTTGCCA TGGGAATTGA CAGAAGTTCC CTTAATTCCC
1351 TTCAAGCTTC ATCCTTTTCC CCGAAGAAGA GGAAGTTTTT CGGTAGTAAG
1401 ACAAGAAAGT CCTTCTTTAT GAGAGGGTCC AAGACGGCCC AAGCCTCAGC
1451 GTCTGATTCA GAGGACGATG CCTCTAAAAA TCCACAGCTC CTTGAGCAGA
1501 CCAAACGACT GTCCCAGAAC TTGCCAGTGG ATCTCTTTGA TGAGCACGTG
1551 GACCCCTCC ACAGGCAGAG AGCGCTGAGC GCTGTCAGTA TCTTAACCAT
1601 CACCATGCAG GAACAAGAAA AATTCCAGGA GCCTTGTTTC CCATGTGGGA
1651 AAAATTTGGC CTCTAAGTAC CTGGTGTGGG ACTGTAGCCC TCAGTGGCTG
1701 TGCATAAAGA AGGTCCTGCG GACCATCATG ACGGATCCCT TTAGTGAGCT
1751 GGCCATCACC ATCTGCATCA TCATCAATAC CGTTTTCTTA GCCGTGGAGC
1801 ACCACAACAT GGATGACAAC TTAAAGACCA TACTGAAAT AGGAACTGG
1851 GTTTTCACGG GAATTTTCAT AGCGGAAATG TGTCTCAAGA TCATCGCGCT
1901 CGACCCTTAC CACTACTTCC GGCACGGCTG GAATGTTTTT GACAGCATCG
1951 TGGCCCTCCT GAGTCTCGCT GATGTGCTCT ACAACACACT GTCTGATAAC
2001 AATAGGTCTT TCTTGGCTTC CCTCAGAGTG CTGAGGGTCT TCAAGTTAGC
2051 CAAATCCTGG CCCACGTAA AACTCTCAT TAAGATCATC GGCCACTCCG
2101 TGGGCGCGCT TGGAAACCTG ACTGTGGTCC TGAATATCGT GGTCTTCATC
2151 TTTTCTGTGG TGGGCATGCG GCTCTTCGGC ACCAAGTTTA ACAAGACCGC
2201 CTACGCCACC CAGGAGCGGC CCAGGCGGCG CTGGCACATG GATAATTTCT
2251 ACCACTCCTT CCTGGTGGTG TTCCGCATCC TCTGTGGGGA ATGGATCGAG
2301 AACATGTGGG GCTGCATGCA GGATATGGAC GGCTCCCCGT TGTGCATCAT

24/27
Figure 5C: SEQ ID NO: 5

2351 TGTCTTTGTC CTGATAATGG TGATCGGGAA GCTTGTGGTG CTTAACCTCT
2401 TCATTGCCTT GCTGCTCAAT TCCTTCAGCA ATGAGGAGAA GGATGGGAGC
2451 CTGGAAGGAG AGACCAGGAA AACCAAAGTG CAGCTAGCCC TGGATCGGTT
2501 CCGCCGGGCC TTCTCCTTCA TGCTGCACGC TCTTCAGAGT TTTTGTGCA
2551 AGAAATGCAG GAGGAAAAAC TCGCCAAAGC CAAAAGAGAC AACAGAAAGC
2601 TTTGCTGGTG AGAATAAAGA CTCAATCCTC CCGGATGCGA GGCCCTGGAA
2651 GGAGTATGAT ACAGACATGG CTTTGTACAC TGGACAGGCC GGGGCTCCGC
2701 TGGCCCCACT CGCAGAGGTA GAGGACGATG TGAATATTG TGGTGAAGGC
2751 GGTGCCCTAC CCACCTCACA ACATAGTGCT GGAGTTCAGG CCGGTGACCT
2801 CCCTCCAGAG ACCAAGCAGC TCACTAGCCC GGATGACCAA GGGGTTGAAA
2851 TGAAGTATT TTCTGAAGAA GATCTGCATT TAAGCATACA GAGTCCTCGA
2901 AAGAAGTCTG ACGCAGTGAG CATGCTCTCG GAATGCAGCA CAATTGACCT
2951 GAATGATATC TTTAGAAATT TACAGAAAAC AGTTTCCCCC AAAAAGCAGC
3001 CAGATAGATG CTTTCCCAAG GGCCTTAGTT GTCACTTTCT ATGCCACAAA
3051 ACAGACAAGA GAAAGTCCCC CTGGGTCCTG TGGTGGAACA TTCGGAAGAC
3101 CTGCTACCAA ATCGTGAAGC ACAGCTGGTT TGAGAGTTTC ATAATCTTTG
3151 TTATTCTGCT GAGCAGTGGA GCGCTGATAT TTGAAGATGT CAATCTCCCC
3201 AGCCGGCCCC AAGTTGAGAA ATTACTAAGG TGTACCGATA ATATTTTCAC
3251 ATTTATTTTC CTCCTGGAAG TGATCCTGAA GTGGGTGGCC TTTGGATTCC
3301 GGAGGTATTT CACCAGTGCC TGGTGCTGGC TTGATTTTCT CATTGTGGTG
2251 GTGTCTGTGC TCAGTCTCAT GAATCTACCA AGCTTGAAGT CCTTCCGGAC
3401 TCTGCGGGCC CTGAGACCTC TCGGGGCGCT GTCCCAGTTT GAAGGAATGA
3451 AGGTTGTCGT CTACGCCCTG ATCAGCGCCA TACCTGCCAT TCTCAATGTC
3501 TTGCTGGTCT GCCTCATTTT CTGGCTCGTA TTTTGTATCT TGGGAGTAAA

25/27
Figure 5D: SEQ ID NO: 5

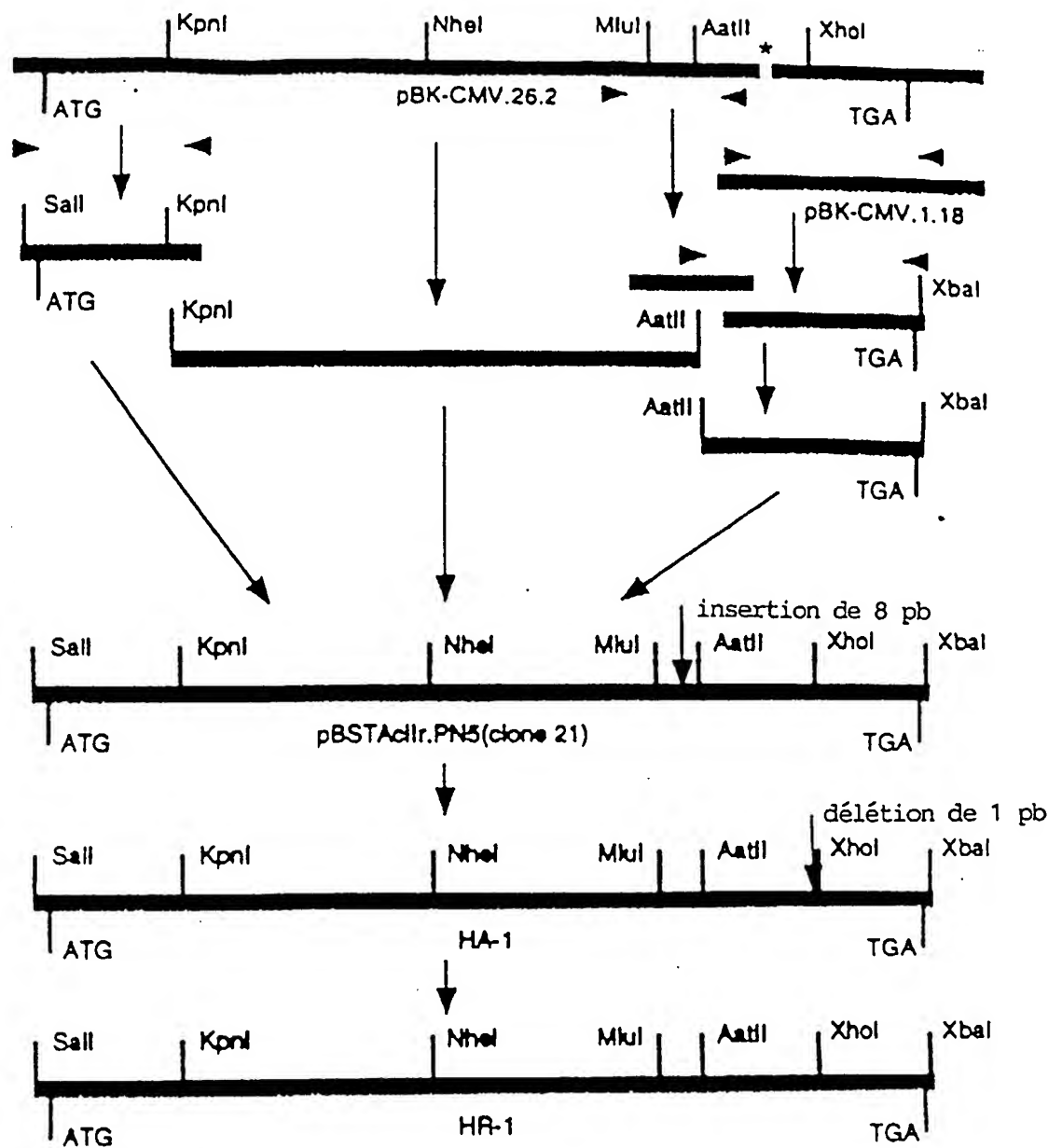
3551 TTTATTTTCT GGAAGTTTG GAAGGTGCAT TAACGGGACA GACATAAATA
3601 TGTATTTGGA TTTTACCGAA GTTCCGAACC GAAGCCAATG TAACATTAGT
3651 AATTACTCGT GGAAGGTCCC GCAGGTCAAC TTTGACAACG TGGGGAATGC
3701 CTATCTCGCC CTGCTGCAAG TGGCAACCTA TAAGGGCTGG CTGGAAATCA
3751 TGAATGCTGC TGTCGATTCC AGAGAGAAAG ACGAGCAGCC GGACTTTGAG
3801 GCGAACCTCT ACGCGTATCT CTACTTTGTG GTTTTTATCA TCTTCGGCTC
3851 CTTCTTTACC CTGAACCTCT TTATCGGTGT TATTATTGAC AACTTCAATC
3901 AGCAGCAGAA AAAGTTAGGT GGCCAAGACA TCTTCATGAC TGAGGAGCAG
3951 AAGAAATATT ACAATGCAAT GAAAAAGTTA GGAACCAAGA AACCTCAAAA
4001 GCCCATCCCA AGGCCCTGA ACAAATGTCA AGCCTTTGTG TTCGACCTGG
4051 TCACAAGCCA GGTCTTTGAC GTCATCATTC TGGGTCTTAT TGTCTTAAAT
4101 ATGATTATCA TGATGGCTGA ATCTGCCGAC CAGCCCAAAG ATGTGAAGAA
4151 AACCTTTGAT ATCCTCAACA TAGCCTTCGT GGTCATCTTT ACCATAGAGT
4201 GTCTCATCAA AGTCTTTGCT TTGAGGCAAC ACTACTTCAC CAATGGCTGG
4251 AACTTATTTG ATTGTGTGGT CGTGGTTCTT TCTATCATTA GTACCCTGGT
4301 TTCCCGCTTG GAGGACAGTG ACATTTCTTT CCCGCCACG CTCTTCAGAG
4351 TCGTCCGCTT GGCTCGGATT GGTCGAATCC TCAGGCTGGT CCGGGCTGCC
4401 CGGGGAATCA GGACCCTCCT CTTTGCTTTG ATGATGTCTC TCCCCTCTCT
4451 CTTCAACATC GGTCTGCTGC TCTTCCTGGT GATGTTTCATT TACGCCATCT
4501 TTGGGATGAG CTGGTTTTCC AAAGTGAAGA AGGGCTCCGG GATCGACGAC
4551 ATCTTCAACT TCGAGACCTT TACGGGCAGC ATGCTGTGCC TCTTCCAGAT
4601 AACCCTTCG GCTGGCTGGG ATACCCTCCT CAACCCCATG CTGGAGGCAA
4651 AAGAACACTG CAACTCCTCC TCCAAGACA GCTGTCAGCA GCCGCAGATA
4701 GCCGTCGTCT ACTTCGTCAG TTACATCATC ATCTCCTTCC TCATCGTGGT

26/27
Figure 5E: SEQ ID NO: 5

4751 CAACATGTAC ATCGCTGTGA TCCTCGAGAA CTTCAACACA GCCACGGAGG
4801 AGAGCGAGGA CCCTCTGGGA GAGGACGACT TTGAAATCTT CTATGAGGTC
4851 TGGGAGAAGT TTGACCCCGA GCGTCGCAG TTCATCCAGT ATTCGGCCCT
4901 CTCTGACTTT GCGGACGCCC TGCCGGAGCC GTTGCCTGTG GCCAAGCCGA
4951 ATAAGTTTCA GTTCTAGTG ATGGACTTGC CCATGGTGAT GGGCGACCGC
5001 CTCCATTGCA TGGATGTTCT CTTTGCTTTC ACTACCAGGG TCCTCGGGGA
5051 CTCCAGCGGC TTGGATACCA TGAAAACCAT GATGGAGGAG AAGTTTATGG
5101 AGGCCAACCC TTTTAAGAAG CTCTACGAGC CCATAGTCAC CACCACCAAG
5151 AGGAAGGAGG AGGAGCAAGG CGCCGCCGTC ATCCAGAGGG CCTACCGGAA
5201 ACACATGGAG AAGATGGTCA AACTGAGGCT GAAGGACAGG TCAAGTTCAT
5251 CGCACCAGGT GTTTTGCAAT GGAGACTTGT CCAGCTTGGA TGTGGCCAAG
5301 GTCAAGGTTC ACAATGACTG AACCCTCATC TAGA

27/27

Figure 6



ATG: codon d'initiation TGA: codon d'arrêt

★
délétion
de 58 pb

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☒ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.